

Zwischen Reputation und Markt: Ziele, Verfahren und Instrumente von (Selbst)Evaluationen außeruniversitärer, öffentlicher Forschungseinrichtungen

Röbbecke, Martina; Simon, Dagmar

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

SSG Sozialwissenschaften, USB Köln

Empfohlene Zitierung / Suggested Citation:

Röbbecke, M., & Simon, D. (1999). *Zwischen Reputation und Markt: Ziele, Verfahren und Instrumente von (Selbst)Evaluationen außeruniversitärer, öffentlicher Forschungseinrichtungen*. (Papers / Wissenschaftszentrum Berlin für Sozialforschung, 99-002). Berlin: Wissenschaftszentrum Berlin für Sozialforschung gGmbH. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-116609>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

P 99 - 002

ZWISCHEN REPUTATION UND MARKT

Ziele, Verfahren und Instrumente von (Selbst)Evaluationen
außeruniversitärer, öffentlicher Forschungseinrichtungen

MARTINA RÖBBECKE ♦ DAGMAR SIMON

Wissenschaftszentrum Berlin für Sozialforschung gGmbH (WZB)
Reichpietschufer 50, D-10785 Berlin

Dr. Martina Röbbecke
Tel. 030 - 254 91 -534
e-mail: roebbecke@medea.wz-berlin.de

Dr. Dagmar Simon
Tel. 030 - 254 91 -588
e-mail: dsimon@medea.wz-berlin.de

Wissenschaftszentrum Berlin für Sozialforschung
Reichpietschufer 50
D-10785 Berlin

Inhaltsverzeichnis

I. Einleitung	S. 7
II. Ausgangspunkte	S. 11
1. Evaluationen im Wissenschaftsbereich	S. 11
2. Politische Rahmenbedingungen	S. 17
3. Konstituierung der Wissenschaftsgemeinschaft G. W. Leibniz und Evaluation durch den Wissenschaftsrat	S. 22
4. Evaluation als Verfahren der Selbstbeobachtung: die Praxis des WZB	S. 25
III. Methodisches Vorgehen	S. 27
IV. Fragestellungen und Zwischenbefunde des Projekts	S. 29
1. Aufgaben, Ziele und Organisation von WGL-Instituten	S. 29
1.1 Aufgaben und Ziele: Entwicklung und Behauptung eines eigenen Profils	S. 33
2. Organisation und Steuerung von WGL-Instituten	S. 39
2.1 Organisation, Planung und Steuerung als Gegenstand der Evaluation	S. 40
3. Die Suche nach angemessenen Indikatoren	S. 51
3.1 Die Dominanz "akademischer" Indikatoren	S. 54
4. Ziele und Verfahren: Beratung und Bewertung	S. 59
4.1 Sein oder Nicht-Sein: Die Evaluation durch den Wissenschaftsrat	S. 65
5. Evaluation als Qualitätssicherung und -förderung	S. 68
5.1 Qualitätssicherung (nicht nur) durch Evaluationen	S. 69
V. Resümee	S. 72
Abkürzungen	S. 76
Literatur	S. 77

Zusammenfassung

In den vergangenen Jahren wurden die wissenschaftspolitischen Diskussionen in erheblichem Maße von der Frage bestimmt, wie - insbesondere angesichts stagnierender oder sogar zurückgehender Ressourcen - die Qualität der Forschung gesichert werden kann. Zunächst konzentrierten sich die Bemühungen zur Qualitätssicherung wesentlich auf die Hochschulen; in jüngster Zeit rücken auch die außeruniversitären Forschungseinrichtungen - insbesondere die zur Zeit durch den Wissenschaftsrat evaluierten Einrichtungen der Blauen Liste - in das Zentrum der Aufmerksamkeit. Das vorliegende Paper setzt sich mit Zielen, Verfahren und Instrumenten von Evaluationen der in ihren Aufgabenstellungen und Zielen ausgesprochen heterogenen Forschungsinstitute der Wissenschaftsgemeinschaft G. W. Leibniz (WGL) auseinander. Im Hinblick auf die Heterogenität gewinnt die Frage nach den adäquaten Bewertungsmaßstäben besondere Relevanz. Dabei geht es zunächst um die Suche nach gültigen Indikatoren - ausgehend von der Annahme, daß die gängigen Wissenschaftsindikatoren das Leistungsspektrum der WGL-Institute - die Verbindung von Grundlagen- und anwendungsorientierter Forschung sowie von Beratungs- und Dienstleistungen - nur unzureichend erfassen. In einem zweiten Schritt wird auf die Ziele von Evaluationen eingegangen und Elemente eines noch weiterzuentwickelnden - auf Qualitätsförderung ausgerichteten - Modells vorgestellt, das interne Selbstbeobachtungs- und externe Begutachtungsprozesse integrieren sollte. Ein weiterer Schwerpunkt der Studie, die Untersuchung der Aufbau- und Ablauforganisation, geht der Hypothese nach, daß die Leistungsfähigkeit einer Forschungseinrichtung entscheidend damit zusammenhängt, welche organisatorischen Lösungen für die Durchführung komplexer Forschungs- und Beratungsaufgaben gefunden werden. Plädiert wird dafür, daß neben den Forschungsergebnissen die Forschungsorganisation - beispielsweise interne Steuerungs- und Managementstrukturen - einen wichtigen Gegenstand von Evaluationen darstellt.

Wer einen Sumpf trockenlegen will, darf nicht die Frösche fragen

Dieter Simon

I. Einleitung

In den letzten Jahren sind Forschungsbewertungen und Evaluierungsmethoden sowohl auf der internationalen wie auf der nationalen Ebene zu wichtigen Themen der wissenschaftspolitischen Diskussionen geworden.

Forschungsbewertungen sind schon heute ein wichtiger Bestandteil des wissenschaftlichen Lebens von außeruniversitären, staatlich finanzierten Forschungseinrichtungen. Diese "Begutachtungen" werden meist von den Beiräten der Forschungseinrichtungen durchgeführt und haben in der Regel einen hohen Stellenwert für die Selbstvergewisserung und die perspektivische Orientierung einer Forschungseinrichtung. Ihre Bedeutung wird zusätzlich verstärkt durch Anforderungen aus der Umwelt der Forschungseinrichtungen nach höherer Transparenz, effizienterem Mitteleinsatz und nach einer Rechenschaftslegung ("accountability") gegenüber Staat und Gesellschaft.

Bei dem vorliegenden Text handelt es sich um erste Auswertungen im Rahmen eines vom Stifterverband für die Deutsche Wissenschaft geförderten Projektes, das sich mit den Zielen, Verfahren und Instrumenten von Evaluationen auseinandersetzt und das einen Beitrag zur Diskussion um die zukünftige Gestaltung der Evaluationen in der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL) leisten will. Es geht zurück auf Erfahrungen des Wissenschaftszentrum Berlin für Sozialforschung (WZB), das schon seit 1994 ein eigenes Verfahren der Selbstbeobachtung konzipiert und erprobt hat. Das Projekt greift Vorschläge des Wissenschaftsrates auf, die Tauglichkeit dieses Verfahrens und der dabei eingesetzten Indikatoren für andere Blaue Liste-Institute zu untersuchen, und es gibt uns zugleich die Möglichkeit einer kritischen Überprüfung der eigenen Ansätze. Eine begrenzte Auswahl der bisherigen Ergebnisse wollen wir im folgenden vorstellen.

In Kapitel II wird zunächst der Kontext der aktuellen Evaluationsdebatten geschildert. Dazu gehören die Diskussionen um Evaluationsverfahren an den Hochschulen, von denen wichtige inhaltliche und konzeptionelle Anregungen auf vergleichbare Debatten an den außeruniversitären Forschungseinrichtungen ausgegangen sind. Gleichwohl unterscheiden sich in den beiden Sektoren die Interessen der beteiligten Akteure, die institutionellen Voraussetzungen und die politischen Rahmenbedingungen. Zunächst werden diese Unterschiede deutlich gemacht, und anschließend wird insbesondere auf die politischen Rahmenbedingungen eingegangen. Zu ihnen gehören die Initiativen von Bund und Ländern, die von allen außeruniversitären Forschungseinrichtungen die Einführung systematischer Evaluationen erwarten, sowie die Aktivitäten des Wissenschaftsrates, der den Zusammenschluß der Einrichtungen der Blauen Liste in der WGL maßgeblich beeinflußt hat. Die Forschungsbewertungen des Wissenschaftsrates prägen die Konturen des zukünftigen Evaluationsverfahrens der WGL. Abgeschlossen wird das Kapitel mit einer kurzen Skizze der Erfahrungen, die das WZB mit seinem Verfahren der Selbstevaluation gewonnen hat.

Die nachfolgenden beiden Kapitel konzentrieren sich auf die zentralen Fragestellungen unseres Projektes und erste Zwischenergebnisse. In Kapitel III werden die Auswahl der Institute und der Gesprächspartner sowie der Aufbau des Interviewleitfadens vorgestellt.

Kapitel IV behandelt zunächst eine Schwierigkeit, die jedes Evaluationsverfahren bewältigen muß, in dessen Rahmen mehrere Institute evaluiert werden: Es müssen Bewertungsmaßstäbe gefunden werden, die Forschungseinrichtungen mit unterschiedlichen Aufgaben und Zielen gerecht werden. Diese Anforderung gilt in besonderem Maße für eine Evaluation der WGL-Institute, die bekanntlich ausgesprochen heterogen sind. Ihr Aufgabenprofil ist durch verschiedene Dimensionen gekennzeichnet: Durch unterschiedliche Gegenstandsbereiche, Forschungstypen und Anwendungskontexte. Die Einrichtungen orientieren sich zum einen an innerwissenschaftlich generierten Relevanzkriterien und an der Anerkennung durch die scientific community, sie sind also auf das spezifische Belohnungssystem der Wissenschaft und den Reputationserwerb verwiesen. Zum anderen müssen die Institute wissenschaftspolitischen Erwartungen gerecht werden, wozu - neben einem Anwendungsbezug und zu erbringenden Beratungs- und Dienstleistungsaufgaben - in jüngerer Zeit auch eine verstärkt geforderte Ausrichtung der Forschungen auf marktfähige Produkte, Verfahren und Dienstleist-

ungen gehört. Die Institute bewegen sich daher in einer Spannung unterschiedlicher Aufgaben und Ziele, die in ihrer ausgeprägtesten Form - "Zwischen Reputation und Markt" - unserer Studie den Titel gegeben hat.

Um das Leistungsspektrum von WGL-Instituten zu qualifizieren, haben wir in fünf Einrichtungen mit verschiedenen Aufgaben und unterschiedlichen disziplinären Orientierungen die Verbindung von Grundlagen- und anwendungsorientierter Forschung sowie die Relevanz der Beratungs- und Dienstleistungsaufgaben untersucht (zu den Instituten und der Zahl der Gespräche vgl. Kapitel III). Darüber hinaus haben wir unsere Annahme überprüft, daß die Aufgabenstellung in erheblicher Weise durch externe Akteure beeinflußt wird und insofern die Auseinandersetzung mit deren Wünschen und Interessen einen eigenen Leistungsbereich der Forschungseinrichtungen darstellt. Ein weiterer Schwerpunkt der Studie, die Untersuchung der Aufbau- und Ablauforganisation, geht der Hypothese nach, daß die Leistungsfähigkeit einer Forschungseinrichtung entscheidend damit zusammenhängt, welche organisatorischen Lösungen für die Durchführung der komplexen Forschungs- und Beratungsarbeiten gefunden werden.

Anschließend folgt ein Abschnitt, der sich aus verschiedenen Perspektiven mit Leistungsindikatoren auseinandersetzt. Zunächst geht es um die Suche nach gültigen Indikatoren für WGL-Institute: Wir gehen der Annahme nach, daß die bisher bekannten Wissenschaftsindikatoren das Leistungsspektrum der WGL-Institute zwischen Grundlagen- und anwendungsorientierter Forschung nicht adäquat erfassen und untersuchen, ob sich die Beratungs- und Dienstleistungen durch Indikatoren bewerten lassen. Die Wahl von angemessenen Indikatoren ist, wie ferner gezeigt werden soll, unter anderem deshalb wichtig, weil von den Indikatoren auch nicht-intendierte Steuerungseffekte ausgehen können.

Der dritte Komplex beschäftigt sich mit den Zielen von Evaluationen. Zunächst gehen wir auf verschiedene Versuche der Kategorisierung von Evaluationen ein; im Vordergrund steht dabei die Unterscheidung zwischen "bewertenden" und "beratenden" Evaluationen. Bei der Begutachtung durch den Wissenschaftsrat stehen die bewertenden Elemente im Vordergrund, und da sein Verfahren auch die zukünftige Evaluation der WGL prägen wird, haben wir ermittelt, wie einzelne Verfahrenselemente von unseren GesprächspartnerInnen beurteilt werden. Abschließend gehen wir auf das wohl meistgenannte Ziel von Evaluationen ein, die "Sicherung der Qualität" der Forschung.

Obwohl von den forschungspolitischen Akteuren immer wieder argumentiert wird, daß Evaluationen der Qualitätssicherung dienen, ist weder klar, was unter "Qualität der Forschung" zu verstehen ist, noch ob und in welcher Form Evaluationen ein geeignetes Instrument zur "Qualitätssicherung" sein können. Daher haben wir unsere GesprächspartnerInnen befragt, was sie unter "guter Forschung" verstehen, was die Ziele von Evaluationen sein können und sollen und welche Maßnahmen aus ihrer Sicht einer "Qualitätssicherung" dienlich sind.

Zuletzt möchten wir den Instituten danken, die sich zur Teilnahme an dieser Studie bereit erklärt haben. Wir sind nicht nur bei der Vorbereitung der Besuche großzügig unterstützt und freundlich aufgenommen worden - darüber hinaus haben sich die einzelnen GesprächspartnerInnen viel Zeit für uns genommen und unsere Fragen mit Geduld, Offenheit und viel Engagement beantwortet. Wir bedanken uns sehr für die anregenden Gespräche.

II. Ausgangspunkte

1. Evaluationen im Wissenschaftsbereich

In beiden staatlich finanzierten Sektoren des bundesdeutschen Wissenschaftssystems, den Hochschulen und den außeruniversitären Forschungseinrichtungen, stehen die Ziele, Instrumente und Verfahren von Evaluationen im Mittelpunkt intensiver Diskussionen.

Beurteilungen durch wissenschaftliche Beiräte sind in vielen außeruniversitären Forschungsinstituten schon seit langer Zeit üblich. Seit dem Anfang der 90er Jahre werden diese internen Verfahren - die hier zukünftig als "Begutachtungen" bezeichnet werden - überarbeitet und beträchtlich erweitert. Offenkundig haben die Zielvorstellungen und konzeptionellen Ansätze der Hochschulen Einfluß auf entsprechende Überlegungen in den Forschungseinrichtungen gehabt, da zentrale Gemeinsamkeiten der Debatten feststellbar sind, die auch im Mittelpunkt dieser Studie stehen werden: Die zur Diskussion stehenden oder bereits implementierten Evaluationsverfahren sind durch eine Verbindung von *internen* Begutachtungen mit *externen* Beurteilungen gekennzeichnet, und sie sollen insbesondere dem Ziel der *Qualitätssicherung* dienen.

Im Unterschied zu den Erörterungen an den außeruniversitären Forschungseinrichtungen setzten vergleichbare Hochschuldebatten bereits Anfang der 80er Jahre ein, und einzelne Reformprojekte - wie etwa der Verbund Norddeutscher Universitäten - haben bereits mehrjährige Erfahrungen vorzuweisen. Diese Ungleichzeitigkeit der Entwicklung spiegelt sich auch in der Literatur wieder: Während die Fülle der Publikationen zu Fragen der Evaluation im Hochschulbereich kaum überschaubar ist, gibt es nur wenige Studien, die sich mit der Evaluation der außeruniversitären Forschungseinrichtungen auseinandersetzen.¹

Wenn trotz der erwähnten Gemeinsamkeiten ein Erfahrungsaustausch zwischen Akteuren aus den Hochschulen und den außeruniversitären Forschungseinrichtungen

¹ Zu diesen seltenen Studien gehören Campbell/Felderer (1997), Krull (1995) und Krull/Winter (1995).

noch nicht recht in Gang gekommen ist, kann das - unter anderem - damit zusammenhängen, daß die institutionellen Besonderheiten, die Interessen und Ziele der beteiligten Akteure sowie die politischen Rahmenbedingungen sehr unterschiedlich sind. Im folgenden sollen in der Form einer knappen Skizze einige Unterschiede zwischen Hochschulen und Forschungseinrichtungen benannt werden. Dieser vergleichend angelegte Überblick wird einige besondere Anforderungen deutlich machen, mit denen die außeruniversitären Forschungseinrichtungen konfrontiert sind. Sie müssen als Kontext der aktuellen Evaluationsdebatten mit bedacht werden; allerdings kann ihnen an dieser Stelle nicht umfassend nachgegangen werden.

Die Evaluationen an den Hochschulen unterscheiden sich zunächst einmal und maßgeblich durch ihren *Gegenstand* von den Evaluationen der Forschungseinrichtungen: Ganz überwiegend konzentrieren sie sich auf die Evaluation von Lehre und Studium, nicht jedoch der Forschung. Zwar wird immer wieder verlangt, die Evaluationen auch auf die Forschung auszudehnen, aber diese Forderung stößt nicht gerade auf die Zustimmung seitens der Gruppe der Professoren, die in großen Teilen schon der Lehrvaluation wenig aufgeschlossen gegenübersteht. Gegenstand der Hochschulevaluationen sind beispielsweise der Aufbau der Studienprogramme, die Verbindung von Studium, Lehre und Praxis, die Ausstattung der Fachbereiche, der Studienverlauf und der Absolventenverbleib. Methodisch liegen den Evaluationen sowohl qualitative Informationen - beispielsweise basierend auf Befragungen von Studierenden und Absolventen - als auch quantitative Daten zugrunde. Mit sogenannten Kennzahlen werden unter anderem die Zahl der Studierenden, die Zahl und Art der Prüfungen sowie Personalbestand und -entwicklung erfaßt. Beide Daten - Umfrageergebnisse und Kennzahlen - gehen in die Bewertungen der Hochschulen ein (vgl. Müller-Böling 1995, HRK 1998 und HRK 1999).

An den außeruniversitären Forschungseinrichtungen dagegen wird die Forschung und somit ein ganz anderer Gegenstandsbereich evaluiert. Die Qualität der Forschung ist bisher ein schwer zu definierendes Konzept geblieben, und noch immer sind in den Evaluationsdebatten die Suche nach Leistungsindikatoren und die Erörterung ihrer Verlässlichkeit und Gültigkeit von zentraler Bedeutung. Neben den Wissenschaftsindikatoren - auf die später eingegangen werden soll - fließen in die Evaluationen ebenfalls strukturelle Daten, etwa zur Personalentwicklung und zur Finanzierung, ein. Darüber hinaus wird in den Evaluationen der Blaue Liste-Institute durch den

Wissenschaftsrat ein dritter Komplex von Informationen erhoben und bewertet. Dabei handelt es sich um Fragen zu Aufbau und Organisationsstruktur, die zum Teil auch wissenschaftspolitisch umstrittene Aspekte berühren: So werden unter anderem die Zahl der Abteilungen, die Aufgabe und die Zusammensetzung der Entscheidungs- und Beratungsgremien, der Anteil an befristeten Stellen sowie die Kooperationen mit Universitäten erfragt. Dieser Ansatz ist in zweierlei Hinsicht interessant: Auf diese Weise werden Evaluationen der Erkenntnis gerecht, daß gute Forschungsleistungen nicht nur das Ergebnis individueller Leistungen, sondern auch Resultat eines betriebsförmig organisierten Prozesses sind, und darüber hinaus wird die angezielte "Qualitätssicherung" insofern ernst genommen, als nicht nur die Qualität der "Produkte" überprüft wird, sondern auch der "Produktionsprozeß" bewertet wird. Gleichwohl bleibt zu fragen, ob es tatsächlich objektive Kriterien für die Bewertung von Organisations- und Entscheidungsstrukturen außeruniversitärer Forschungseinrichtungen gibt und ob eine solche Bewertung überhaupt Bestandteil von Evaluationen sein sollte.

Ein weiterer zentraler Unterschied zwischen Hochschulen und außeruniversitären Forschungseinrichtungen besteht darin, daß in den Hochschulen interne *Akteure*, die damit auch die innere *Reform* der Universitäten befördern wollten, die wohl wichtigsten Beiträge zur Durchsetzung von Evaluationen geleistet haben. Zwar haben auch die Hochschulbewertungen und "Rankings" von verschiedenen großen Zeitschriften zur kritischen Diskussion über Methoden und Ziele solcher Bewertungen beigetragen und, nicht zuletzt, erhebliche Qualitätsunterschiede zwischen den Hochschulen deutlich gemacht. Infolgedessen sahen sich die Hochschulen von seiten der Medien, der interessierten Öffentlichkeit und der staatlichen Ministerialbürokratie einem wachsenden Reformdruck ausgesetzt. Jedoch ist die Bedeutung nicht zu unterschätzen, die die studentischen Lehrevaluationen bzw. das originäre Interesse der Studierenden an der Verbesserung der Lern- und Studienbedingungen für die Akzeptanz und Durchführung von Evaluierungen gehabt haben. Darüber hinaus haben auch ProfessorInnen und Hochschulleitungen die Einführung von Evaluationen an den Hochschulen als ein Reformprojekt begriffen und bescheinigen den Evaluationen heute, daß sie nicht nur zur Verbesserung der Studienbedingungen, sondern auch zu veränderten Kommunikationsstrukturen zwischen Studierenden und Lehrenden beitragen: "Verfahren, in denen Selbstreflexion und externe Begutachtung kombiniert werden, sind dazu geeignet, die Kommunikation zwischen Lehrenden einerseits, Studierenden andererseits und

zwischen beiden Gruppen zu öffnen für die Probleme in Lehre und Studium" (Lüthje 1995, S. 158).

Die aktuellen Evaluationsdebatten an den außeruniversitären Forschungseinrichtungen sind dagegen erheblich durch externe (politische) Forderungen veranlaßt worden, was für die internen Akteure nicht unbedingt ein Anlaß zu Reformhoffnungen gewesen sein mag. Man wird allerdings zwischen den internen und den externen Forschungsbewertungen differenzieren müssen: Für viele Forschungseinrichtungen sind interne Beurteilungen durch die Beiräte - für die sich bezeichnenderweise der Begriff "Begutachtung" gehalten hat - ein selbstverständlich gewordener Bestandteil des wissenschaftlichen Lebens. Die meisten Forschungseinrichtungen kennen beispielsweise Begutachtungen im Rahmen von Neuberufungen, und viele Einrichtungen oder Institutsabteilungen werden in mehr oder weniger regelmäßigen Abständen von den wissenschaftlichen Beiräten begutachtet. Häufig sind die Beiräte unterstützend tätig; so ist etwa nach Auffassung der Helmholtz-Gemeinschaft Deutscher Forschungszentren "... der wichtigste Effekt, der durch diese Beiräte ausgeübt wird, die Beratung des Managements bei der laufenden wissenschaftlichen Steuerung des Gesamtzentrums oder des jeweiligen Instituts." (HGF 1997, S. 30). Zwar richten sich gegen diese Begutachtungen bekannte Einwände: Die Beiräte seien, so wird von Beobachtern berichtet, bereits erheblich in die Verantwortung für "ihre" Einrichtung einbezogen und daher in ihren Beurteilungen nicht mehr unabhängig. Jedoch versuchen viele Forschungseinrichtungen im eigenen Interesse, diese Mängel - die als Schwächen vieler Peer Review-Verfahren gelten - durch eine Verbesserung des Beiratssystems zu beheben. Beispielsweise hat die Max-Planck-Gesellschaft (MPG) für die Auswahl, Zusammensetzung und Verweildauer der wissenschaftlichen Beiräte differenzierte "Regelungen für Fachbeiräte" geschaffen (vgl. MPG 1997).

Die Erfahrungen mit externen Evaluierungen und daher auch die Vorstellungen und Erwartungen an zukünftige Evaluationsverfahren fallen dagegen ambivalenter aus. Die Einrichtungen der Blauen Liste werden seit den 70er Jahren durch den Wissenschaftsrat beurteilt, und auf Wunsch des BMBF hat der Wissenschaftsrat bisher drei sogenannte "Querschnittsbewertungen" auf den Gebieten der Umweltforschung, der Materialforschung und der Energieforschung durchgeführt, bei denen neben den entsprechenden Aktivitäten aller einschlägig arbeitenden außeruniversitären Forschungseinrichtungen - von den Großforschungszentren über die Max-Planck-Gesellschaft, die Fraunhofer-

Gesellschaft und die Blaue Liste-Institute bis zu den Ressortforschungseinrichtungen des Bundes und der Länder - auch die Forschung und Lehre an den Universitäten einbezogen wurde. Bei aller Kritik im Detail stießen diese Bewertungen, vermutlich nicht zuletzt aufgrund der Reputation des Wissenschaftsrates, auf weitgehende Akzeptanz seitens der Forschungseinrichtungen.

Anders verhielt es sich dagegen bei externen Evaluierungen, die offenkundig von forschungs- bzw. industriepolitischen Interessen dominiert waren. Beispielsweise wurden die Begutachtungen des Forschungszentrums Jülich und des Forschungszentrums Karlsruhe durch die sogenannte "Weule-Kommission" (1993/94) von beiden Zentren scharf kritisiert. Diese Kommission, die aus Repräsentanten führender Industrieunternehmen bestand, gelangte zu den bemerkenswerten und nicht ganz unerwarteten Empfehlungen, den Anteil der industrierelevanten Forschungen massiv zu erhöhen, die Themenfelder gemeinsam festzulegen und Planungs- und Projektentscheidungen im Einvernehmen mit Industrievertretern zu treffen (vgl. Röbbcke 1997, S. 151 f.).

Die Vorstellungen über externe Evaluationen - die Erwartungen, aber auch die Befürchtungen - dürften maßgeblich durch die Evaluation der außeruniversitären Forschungslandschaft der DDR geprägt worden sein. Im Rahmen seiner Forschungsbewertungen beurteilte der Wissenschaftsrat erstmals nicht nur einzelne Institute oder Forschungsfelder, sondern mehrere Forschungseinrichtungen in Gänze (s.u.). Nach den damit verbundenen Erfahrungen und Folgen - insbesondere den Schließungen von Instituten, die sich im Rahmen der laufenden Evaluation der Blauen Liste durch den Wissenschaftsrat wiederholen - erscheinen Evaluationen weniger als ein internes Reforminstrument denn als ein extern dominiertes und politisch beeinflusstes Kontroll- und Steuerungsverfahren mit zum Teil gravierenden Konsequenzen.

Als letzter wichtiger Unterschied soll auf die unterschiedlichen *politischen Rahmenbedingungen* aufmerksam gemacht werden. Bekanntlich wurden die Hochschulen in den vergangenen Jahren massiv kritisiert: Zu lange Studiendauer, hohe Abbruchquoten und ein im internationalen Vergleich hohes Alter der Absolventen führten zu erheblichen Zweifeln an ihrer Leistungsfähigkeit. Während den Hochschulen verkrustete Strukturen und fehlendes Engagement des Lehrpersonals vorgeworfen wurde, wiesen die Hochschulen ihrerseits die Vorwürfe mit Hinweis auf die chronische Unterfinanzierung und bürokratische Reglementierungen durch den Staat zurück. Mit der Hilfe von

Evaluationen scheint eine Behebung von Legitimationsproblemen und eine Neugestaltung des Verhältnisses zwischen Staat und Hochschulen möglich zu sein: Durch einen Ausweis ihrer Leistungsfähigkeit soll die gesellschaftliche Akzeptanz verbessert werden - Transparenz und Rechenschaftslegung ("accountability") sind hier die Schlagworte. Im Rahmen von Evaluationen sollen aber nicht nur die Leistungen der Hochschulen belegt, sondern zugleich Verbesserungsmöglichkeiten ("improvement") von Lehre und Studium identifiziert werden. Im Gegenzug erwarten die Hochschulen eine höhere Autonomie, also eine partielle Befreiung von bürokratischen Regulierungen, die eine Voraussetzung für die Umsetzung von Gestaltungsoptionen darstellt und ihnen - etwa in der Form von Globalhaushalten oder "Experimentierklauseln" - zugesichert wurde.

Ob und inwieweit sich diese hohen Erwartungen an Evaluationen erfüllen werden, sei zunächst dahingestellt. In den Niederlanden, deren Evaluationsverfahren in mehrerer Hinsicht das Vorbild für bundesdeutsche Ansätze ist, werden "improvement" und "accountability" als zwei entgegengesetzte Pole betrachtet, zwischen denen sich die Verfahren der Qualitätssicherung bewegen: "It will not always be easy to combine improvement and accountability. Being responsible for the system of Extern Quality Assessment is like navigating between Scylla and Charybdis."² In ähnlicher Form sind diese Gegensätze auch Merkmale der außeruniversitären Forschungsevaluationen, darauf soll an anderer Stelle näher eingegangen werden.

Unter vergleichenden Aspekten ist bemerkenswert, daß "Rechenschaftslegung" und "Transparenz" auch von den außeruniversitären Forschungseinrichtungen verlangt werden. Die Rhetorik ähnelt sich zwar teilweise, aber die Forderung der Bundesregierung nach regelmäßigen Evaluationen wird in einem Kontext erhoben, der - im Unterschied zu den Hochschulen, denen eine höhere Autonomie zugesagt wird - durch eine ausgeprägtere politische Einflußnahme auf Forschungsthemen und -inhalte der außeruniversitären Forschungseinrichtungen gekennzeichnet ist. Auf die komplexen politischen Rahmenbedingungen soll im nächsten Kapitel ausführlich eingegangen werden.

² Vroeijenstijn 1995: XIV. Er erläutert sein schönes Bild wie folgt: "When one aims only at improvement, the system will be shipwrecked against the "Scylla" because the outside stakeholders will ask for accountability and design their own External Quality Assessment system. If accountability is emphasised too much, the system will disappear in the "Charybdis", because improvement will be hindered or even made impossible. The challenge is to keep on course and, by doing so, reconcile the two purposes in one system." (S. 33)

2. Politische Rahmenbedingungen

Die aktuellen Evaluationsdebatten sind in hohem Maße durch wissenschafts- und forschungspolitische Rahmenbedingungen geprägt. Vor allem zwei Akteure sind hervorzuheben: der Wissenschaftsrat und das Bundesministerium für Bildung und Forschung (BMBF).³

Der Wissenschaftsrat hat sich in dem letzten Jahrzehnt zu einem einflußreichen wissenschaftspolitischen Akteur mit hoher Reputation entwickelt, dessen Bedeutungszuwachs eng mit seiner Evaluationstätigkeit zusammenhängt. Er ist 1957 als wissenschafts- und forschungspolitische Beratungseinrichtung gegründet worden und hat bis zum Ende der 80er Jahre auf Wunsch der Bundesregierung einzelne Hochschulen oder Forschungseinrichtungen begutachtet. Außerdem wurde der Wissenschaftsrat von der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) gebeten, die Überprüfung der Blaue Liste-Institute zu übernehmen. Diese Einrichtungen, auf deren gemeinsame Finanzierung sich Bund und Länder in der 1975 abgeschlossenen "Rahmenvereinbarung Forschungsförderung" verständigt hatten, sollen regelmäßig daraufhin bewertet werden, ob sie die Voraussetzungen der gemeinsamen Förderung weiterhin erfüllen, also die "überregionale Bedeutung" und das "gesamtstaatliche wissenschaftspolitische Interesse".⁴ Seit 1979 hat der Wissenschaftsrat diese Institutsbegutachtungen durchgeführt.⁵

Gleichwohl gehörten Forschungsbewertungen nicht zu den wichtigsten Aufgaben des Wissenschaftsrates. Das sollte sich nach 1990 als Folge der Deutschen Einheit grundlegend ändern: Der Wissenschaftsrat hatte die Aufgabe, die außeruniversitären Forschungseinrichtungen der DDR, also die Akademie der Wissenschaften (AdW), die Akademie der Landbauwissenschaften und die Bauakademie zu evaluieren. Dabei ging es nicht nur um die Frage, welche Forschungsleistungen positiv zu bewerten waren, sondern der Wissenschaftsrat machte zugleich Vorschläge, in welchem institutionellen Kontext die positiv bewerteten Forschergruppen weiter tätig sein sollten. Insofern nahm

³ Das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie wurde im Oktober 1998 umbenannt.

⁴ Vgl. Ausführungsvereinbarung zur Rahmenvereinbarung Forschungsförderung über die gemeinsame Förderung von Einrichtungen der wissenschaftlichen Forschung, § 1 Abs. 5

⁵ Vgl. Maurer (1995)

der Wissenschaftsrat großen Einfluß auf die sogenannte "Neuordnung" der Forschungslandschaft; unter anderem empfahl er, welche Institute erhalten oder aufgelöst werden sollten, er wirkte auf Neugründungen ein und betrieb mit besonderem Nachdruck die (dann kaum erfolgreiche) Integration der AdW-WissenschaftlerInnen in die Hochschulen.⁶

Der Wissenschaftsrat ist, ohne daß er es je intendiert hätte, sozusagen nebenbei zu einer "Evaluationsagentur"⁷ geworden. Bei aller Kritik an den Ergebnissen der Akademie-Evaluationen wurde ihm zugestanden, daß er Verfahren und Instrumente entwickelt habe, mit denen eine schwierige Aufgabe in außerordentlich kurzer Zeit erfolgreich bewältigt werden konnte. Diese Verfahrenselemente prägen, in leicht modifizierter Form, ebenfalls die erneute Evaluation aller Blaue Liste-Einrichtungen durch den Wissenschaftsrat.

Daneben haben die Evaluationen der außeruniversitären Forschungseinrichtungen in den neuen Bundesländern aber auch dazu geführt, daß nicht nur die Blaue Liste-Institute, sondern auch die anderen Forschungseinrichtungen in den alten Ländern einem zunehmenden Legitimationsdruck ausgesetzt wurden. Zunehmend wurde kritisiert, daß diese Forschungsinstitute sich nur selten einem auch nur vergleichbar strengen Prüfverfahren stellen mußten. Der Druck auf die Forschungseinrichtungen wurde seitens der Universitäten noch verstärkt, die - ob in den alten oder den neuen Bundesländern - befürchten mußten, daß ein weiteres Anwachsen des außeruniversitären Forschungssektors durch die Folgeeinrichtungen der Akademien der DDR zu ihren Lasten gehen werde. In den Universitäten der alten Bundesländer wird schon seit Jahren über eine drohende "Auswanderung" der Forschung diskutiert und unter dem Stichwort einer "Reorganisation der außeruniversitären Forschung" deren partielle Zurückführung in die Hochschulen verlangt, und auch die Hochschulen der neuen Länder befürchten angesichts des in vielen Fächern notwendigen Auf- und Ausbaus der Forschung eine potentielle Konkurrenz durch die ehemaligen Akademie-Einrichtungen.

⁶ Vgl. Mayntz (1994) und Wolf (1996)

⁷ Dieter Simon, der ehemalige Vorsitzende des Wissenschaftsrates, hält die "beharrliche Gesamtevaluation der Wissenschaft in Deutschland" für eine der wichtigsten Unternehmungen des Wissenschaftsrates. Er erkennt allerdings auch nicht die Furcht der Mitglieder, daß diese "keineswegs berauschende Daueraufgabe" den Wissenschaftsrat zur "Evaluationsmaschine" mache (vgl. Simon 1997).

Vermutlich haben diese Diskussionen nicht unerheblich zu einem politischen Klima beigetragen, in dem Bund und Länder die außeruniversitären, staatlich finanzierten Forschungseinrichtungen erfolgreich dazu auffordern konnten, die internen Begutachtungsverfahren auszubauen und externe Evaluationen zu veranlassen bzw. zu akzeptieren. Die Ziele und Hintergründe dieses folgenreichen Verlangens sollen kurz skizziert werden.

Der internationale ökonomische Wettbewerb der letzten Jahre hat in nahezu allen westlichen Industriestaaten direkte Auswirkungen auf die unterschiedlichen Wissenschaftssysteme gehabt. Aufgrund der Annahme, daß für den Erhalt führender Positionen und für die Sicherung zukünftiger Märkte die wissensbasierten Technologien eine ausschlaggebende Rolle spielen werden, sind "internationale Wettbewerbsfähigkeit" und "Innovationsorientierung" zu zentralen Begriffen der forschungspolitischen Diskussionen geworden.

Auch in der Bundesrepublik Deutschland ist diese Entwicklung feststellbar: So setzt die Bundesregierung deutliche forschungspolitische Akzente im Bereich der sogenannten "Zukunftstechnologien", und gleichzeitig wird, insbesondere angesichts sinkender Ressourcen für Forschung und Technologie, von den staatlich finanzierten Forschungseinrichtungen erwartet, daß sie einen öffentlichen Nachweis qualitativ hochstehender, effizienter und wirtschaftlicher Leistungen erbringen.

Die Bundesregierung setzt also auf eine "Doppelstrategie", auf "Innovation durch mehr Flexibilität und Wettbewerb" - so lautet der programmatische Titel der "Leitlinien", die das BMBF 1996 veröffentlichte. Darin begründet das BMBF die forschungspolitische Prioritätensetzung auf Innovationsorientierung; es kündigte ferner eine Flexibilisierung der Mittelvergabe und administrative Erleichterungen an, machte jedoch gleichzeitig deutlich, daß von den Forschungseinrichtungen ein stärkerer Wettbewerb um die gleichbleibend knappen staatlichen Ressourcen erwartet wird.

In einem Dokument der Bund-Länder-Kommission mit dem Titel "Sicherung der Qualität der Forschung" wurden 1997 diese Überlegungen präzisiert. Zunächst einmal wurde an einen Beschluß der Regierungschefs von Bund und Ländern vom 18.12.96 erinnert, in dem diese eine Evaluation aller gemeinsam geförderten Forschungseinrichtungen verlangten. Darüber hinaus machte die BLK deutlich, daß sie über dieses einmalige

Vorhaben hinaus von allen Forschungseinrichtungen erwartet, zukünftig eigenständige und kontinuierliche Maßnahmen zur Qualitätssicherung einzuführen. Dazu zählt sie erstens den bereits erwähnten Zusammenhang von Flexibilisierung und Wettbewerbsorientierung, zweitens die Transparenz und Kontrolle des effizienten Mitteleinsatzes und drittens schließlich die Bewertung der wissenschaftlichen Leistungen. Konkret wird daraus ein "Gesamtpaket", das die Einführung von Globalhaushalt und Programmbudget, eine Kosten-Leistungs-Rechnung und schließlich auch die hier interessierende Evaluation beinhaltet.

Es zeichnet sich bereits ab, daß die Forschungseinrichtungen unterschiedliche Strategien einsetzen, um sich den neuen Herausforderungen zu stellen. Welche Strategie eine Einrichtung wählt, hängt unter anderem damit zusammen, wie die internen Beurteilungsverfahren bisher beschaffen sind und, nicht zuletzt, wie das politische Durchsetzungsvermögen gestaltet ist. So ist es der *Max-Planck-Gesellschaft* gelungen, die Forderung nach einer externen Evaluierung der Inhalte und Leistungen ihrer einzelnen Institute weitgehend abzuwehren. Sie hat sich mit den Zuwendungsgebern auf eine sogenannte "systemische Evaluation" verständigt, in deren Rahmen die zentralen Struktur- und Organisationsmerkmale der MPG untersucht und bewertet werden. Eine international zusammengesetzte Expertenkommission soll die beiden großen Förder- und Trägerorganisation der Grundlagenforschung gemeinsam evaluieren: "DFG und MPG haben dazu einen gemeinsamen Vorschlag vorgelegt, dessen Ziel die Bewertung der Stellung der beiden Organisationen in der deutschen Forschungslandschaft ist. Es geht dabei nicht um einzelne Projekte oder Institute, sondern um die Frage, ob die Prinzipien, Verfahren und Instrumente von DFG und MPG heutigen Anforderungen von Wissenschaft, Forschung und Entwicklung gerecht werden." (BMBF 1998, S. 88).

Zwar kommt auch die MPG nicht umhin, regelmäßige externe Forschungsbewertungen der Institute durchzuführen. Dabei kann sie sich jedoch auf die bestehenden Fachbeiräte stützen, die schon bisher nicht nur beratende, sondern auch bewertende Aufgaben haben. Die Fachbeiräte lassen dem Präsidenten der MPG alle zwei Jahre einen differenzierten Bericht zur Beurteilung des jeweiligen Instituts zukommen. Die MPG will dieses Fachbeiratssystem fortentwickeln und die Institute zu verschiedenen Forschungsfeldern zusammenfassen, um sie in einem sechsjährigen Turnus unter vergleichenden Gesichtspunkten beurteilen lassen zu können. Dazu sollen die Fachbeiräte um

zwei externe Wissenschaftler ergänzt werden, die an allen Begehungen der in einem Forschungsfeld zusammengefaßten Institute teilnehmen.

Aus der Sicht der MPG dienen die externen Forschungsbewertungen neben der Beratung der Institute insbesondere der zentralen Ressourcenallokation: "Die Fachbeiratsberichte sind wesentliche Quelle der institutsbezogenen Entwicklungsplanung der Gesellschaft... Die damit gegebenen Steuerungsmechanismen ermöglichen in einem begrenzten Umfang eine differenzierte Reaktion auf Entwicklungen in den Instituten und schaffen für die Max-Planck-Gesellschaft eine gewisse Flexibilität bei der Steuerung der Ressourcenausstattung ihrer Forschungseinrichtungen" (MPG 1997, S. 42 f.). Die Fortentwicklung der Evaluationen zu einem Steuerungsinstrument der Ressourcenallokation setzt allerdings einen spezifischen Finanzierungsmodus voraus: Die MPG handelt den Gesamthaushalt mit den Zuwendungsgebern aus und leitet sie - bis auf diejenigen Mittel, die zur Finanzierung der zentralen Aufgaben notwendig sind - an ihre Institute weiter.⁸

Die Ausgangssituation der *Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz* (WGL) ist dagegen deutlich schwieriger. Sie besitzt die Möglichkeit einer differenzierten Ressourcenallokation, die einen Anreiz für besonders positiv evaluierte Institute darstellen könnte, bisher nicht, da die WGL den Charakter eines Dachverbandes ohne zentral zu vergebende Finanzmittel hat. Um die Herausforderungen, mit denen die WGL derzeit konfrontiert ist, etwas deutlicher zu machen, soll im folgenden zunächst kurz auf die Aufgaben ihrer Institute, die historische Entwicklung und die Struktur der WGL eingegangen werden.

⁸ Auf die *Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren* (HGF) und die *Fraunhofer-Gesellschaft* (FhG) kann an dieser Stelle nur am Rande eingegangen werden. Der Senat der HGF hat 1997 Grundsätze und Verfahren der Begutachtungen in der Helmholtz-Gemeinschaft verabschiedet, die insbesondere für die Begutachtung der jeweiligen Zentren durch externe Gremien gelten sollen. Dabei weist er ausdrücklich darauf hin, daß die Beurteilung der einzelnen Zentren durch interne und externe Gremien bereits seit Jahren praktiziert wird; neu ist dagegen der zentrenübergreifende "Strategiefonds" (vgl. HGF 1997). Die FhG, deren Erfolg als Trägerorganisation der anwendungsorientierten Forschung sich am Markt erweist, hat - ähnlich wie die MPG - eine "Systemevaluation" erfahren. Unter anderem hat die Evaluierungskommission der FhG empfohlen, die Beurteilungsverfahren durch Instituts-Audits und Kundenbefragungen auszubauen (vgl. Bericht der Evaluierungskommission, November 1998).

3. Konstituierung der Wissenschaftsgemeinschaft G. W. Leibniz und Evaluation durch den Wissenschaftsrat

Die Forschungs- und Serviceeinrichtungen der Blauen Liste sind unter unterschiedlichen historischen Bedingungen gegründet und in ihrer Aufgabenstellung weiterentwickelt worden. Ihnen fehlt - im Vergleich zu der Max-Planck-Gesellschaft und der Fraunhofer-Gesellschaft - ein verbindendes Element wie eine gemeinsame Forschungsorientierung, wissenschaftliche Ausrichtung oder wissenschaftspolitische Aufgabenstellung, das mehr als das gemeinsame Band der Finanzierung durch Bund und Länder darstellen würde. Es handelt sich um diejenigen Institute, die 1975 in der "Rahmenvereinbarung Forschungsförderung" über die gemeinsam von Bund und Ländern getragene Förderung von Forschung aufgeführt sind. Vorläufer dieser Rahmenvereinbarung war das "Königsteiner Staatsabkommen" zwischen den Bundesländern, das sich neben der Max-Planck-Gesellschaft, der Deutschen Forschungsgemeinschaft auch auf diese Institute bezog, "die jeweils von einem bestimmten Bundesland getragen wurden, deren Forschungsthemen jedoch nicht nur eine auf dieses beschränkte regionale Bedeutung hatten" (Hohn/Schimank 1990, S. 136). Bei der Schaffung der gemeinsamen Rahmenvereinbarung zeigte sich hinsichtlich der Festlegung, welche Institute vom "Typus Königsteiner Abkommen" aufzunehmen seien und fortan als "Blaue Liste" fungieren sollten, schon die grundsätzliche Problematik der heterogenen Zusammensetzung. Zudem waren Akteure mit zum Teil differierenden Interessen im Spiel, insbesondere die Bundes- und Landesministerien. Ursprünglich wurden bei den Beratungen über 300 Forschungseinrichtungen von unterschiedlichen Seiten ins Spiel gebracht; bei der Verabschiedung der Rahmenvereinbarung Forschungsförderung, die einen Kompromiß der verschiedenen Akteursinteressen darstellte, wurden schließlich zwei Drittel der "Königsteiner Institute" und 22 weitere Einrichtungen integriert.

Besondere Aufmerksamkeit erlangte die heterogene Zusammensetzung der Blaue Liste-Institute nach der laut Einigungsvertrag vorgesehenen Einpassung der aus der DDR übernommenen Forschungseinrichtungen in die gemeinsame Forschungsstruktur, die eine Erhöhung der Zahl der Blauen Liste-Institute von 48 auf rd. 80 zur Folge hatte. Dadurch erreichten die Blaue Liste-Institute ein Finanzierungsvolumen in der Höhe der Max-Planck-Gesellschaft oder der Deutschen Forschungsgemeinschaft. Es gab schon Ende der achtziger Jahre Versuche, sich in der Blauen Liste (ein

“forschungspolitisches Aggregat”, vgl. ebenda, S. 135 ff). über gemeinsame Orientierungen, Aufgabenstellungen, Interessen und Interessenvertretung zu verständigen: 1989 wurde die Arbeitsgemeinschaft Blaue Liste als freiwilliger, loser Zusammenschluß von Instituten mit dem Ziel gegründet, einen Erfahrungsaustausch untereinander zu organisieren und kollektiv in der wissenschaftspolitischen Öffentlichkeit aufzutreten, die gegenüber der Konstituierung einer "vierten Säule" in der deutschen Forschungslandschaft eher kritisch eingestellt war. Jedoch erwiesen sich die zahlenmäßige Erweiterung und insbesondere die Empfehlungen des Wissenschaftsrats zur Neuordnung der Blauen Liste als wichtige Promotoren für den Aufbau einer Wissenschaftsorganisation und der entsprechenden Strukturen.

Für die Institute der Blauen Liste spielten Qualitätskontrolle und wissenschaftliche Leistungsbewertung schon immer eine Rolle. Neben Evaluierungen durch den Wissenschaftsrat wurden zum Teil Kontrollen durch wissenschaftliche Beiräte und Evaluationsgremien ausgeübt, die Evaluierung der Einrichtungen in ihrer Gesamtheit stellt jedoch ein Novum dar. "Im vereinten Deutschland wird die Blaue Liste eine Größenordnung erreichen, bei der auf eine die einzelnen Forschungseinrichtungen *übergreifende Anpassung* an die Entwicklung des Forschungsbedarfs ebensowenig verzichtet werden kann wie bei Einrichtungen der MPG, der FhG oder bei Großforschungseinrichtungen" (Wissenschaftsrat 1993, S. 30).

In der Stellungnahme des Wissenschaftsrats wurden für alle Institute Maßnahmen zur Qualitätssicherung mit unterschiedlichen Elementen (u.a. interne Kontrollen durch die wissenschaftlichen Institutsbeiräte und externe Evaluationen) empfohlen und gemeinsame Charakteristika der Blaue Liste-Institute herausgestellt: So ihre Bedeutung als flexibles Element der wissenschaftspolitischen Steuerung der Länder und des Bundes. Als Instrument der föderalen Forschungspolitik können sie Aufgaben übernehmen, die weder von den Hochschulen noch von anderen außeruniversitären Einrichtungen wahrgenommen werden können (vgl. Wissenschaftsrat 1993, S. 20 ff.).

Im Kontext der wissenschaftspolitischen Debatte über Profilbildung und Qualitätssicherung, die durch die Empfehlungen des Wissenschaftsrats einen zusätzlichen Input erhielt, führte der Konstituierungsprozeß von der AG Blaue Liste über die Wissenschaftsgemeinschaft Blaue Liste zur Wissenschaftsgemeinschaft G. W. Leibniz und damit zur Etablierung formaler Strukturen einer Dachorganisation (Präsidium,

Sektionen, Ausschüsse, seit jüngstem Senatsbildung). In der seit 1993 intensiv geführten Selbstverständigungsdiskussion über ein gemeinsames Profil und gemeinsame Aufgabenstellung, strukturelle Rahmenbedingungen und Qualitätssicherung der Forschung finden sich entscheidende Elemente der Wissenschaftsratsempfehlungen wieder. Die Vielfalt der Institute wurde als Stärke herausgestellt, die Wissenschaftsgemeinschaft als ein Modell zur Verwirklichung einer kooperativen, flexiblen Wissenschaftslandschaft mit dezentraler Verantwortung, Kompetenz und themenangepaßter Reaktionsfähigkeit empfohlen. Intern wurde die Entwicklung einer gemeinsamen Kooperationskultur hinsichtlich der (transdisziplinären) Bearbeitung von Forschungsfeldern mit entsprechenden Synergieeffekten angestrebt.

Eine herausragende Bedeutung nahm und nimmt in dem Selbstverständigungsprozeß die Frage von Evaluationen ein, insbesondere vor dem Hintergrund, daß der Wissenschaftsrat definitiv erklärt hatte, die Evaluierung der Blaue Liste-Institute in ihrer Gesamtheit nicht noch ein weiteres Mal vorzunehmen. Weitere Überlegungen führten dazu, diese Aufgabe dann der WGL selber zu übertragen. In den Vorstellungen der WGL über Evaluationen sollten Vergleiche von Forschungsinstituten nur unter Berücksichtigung von Struktur und Aufgabenstellung vorgenommen, dementsprechend die Bewertungskriterien der Vielfalt und Komplexität der Blauen Liste gerecht werden und die qualitative Bewertung sich an der Erfüllung des jeweiligen wissenschafts- und gesellschaftspolitischen Auftrags und der Bedeutung der Arbeit für die angesprochenen Zielgruppen orientieren. Die Diskussion über Ziele und Zwecke von Forschungsevaluierungen, unterschiedliche Modelle und Methoden wurde dagegen eher nur am Rande geführt.

Durch den einheitlichen, überwiegend akademisch geprägten Kriteriensatz des Wissenschaftsrats zur Bewertung der Blaue Liste-Institute erlangte die Frage nach (spezifischen) Bewertungskriterien sowohl in den laufenden Evaluationen durch den Wissenschaftsrat eine gewisse Bedeutung als auch in den internen Diskussionen der WGL-Gremien über zukünftige Bewertungskriterien bei dem noch zu konzipierenden zukünftigen Evaluationsmodell.

So gesehen hat der Prozeß der Evaluierung durch den Wissenschaftsrat auf die Blaue Liste-Institute insgesamt eine integrierende Funktion ausgeübt in dem Sinn, daß sie eine (neue) Rolle als kollektiver Akteur im bundesdeutschen Wissenschafts- und Forschungssystem einzunehmen beginnen.

4. Evaluation als Verfahren der Selbstbeobachtung: die Praxis des WZB

Der in den letzten Jahren zunehmend enger gewordene Ressourcenrahmen für Wissenschaft und Forschung sowie der erhöhte Außen- und Legitimationsdruck insbesondere für außeruniversitäre Institute stellen die Forschungseinrichtungen vor neue Herausforderungen, die insbesondere ihre Flexibilität und damit interne Steuerungsfähigkeit tangieren. Eine kontinuierliche Überprüfung und Verbesserung von internen Verfahren und Kommunikationsprozessen im Sinne von qualitätssichernden Maßnahmen sollte eigentlich zum "Alltagsgeschäft" von Forschungsinstituten gehören, um eine kritische Reflexion ihrer Leistungsfähigkeit und zukünftiger Herausforderungen zu gewährleisten, was jedoch nicht unbedingt der Fall ist. Die veränderten Rahmenbedingungen verdeutlichen darüber hinaus die Notwendigkeit von internen Verfahren zur Überprüfung und Förderung der Leistungen der Forschung, ihrer Bedingungen und Voraussetzungen.

Institutionelle Selbstbeobachtungsverfahren sollen dazu beitragen, den Leistungsstand der jeweiligen Einrichtung zu erfassen, Stärken und Schwächen zu analysieren und Perspektiven zukünftiger Entwicklung zu entfalten. Der Institutionalisierung eines solchen Verfahrens liegt primär ein Verständnis eines kollektiven Lernprozesses und weniger einer Kontrollroutine zugrunde.

Im WZB wurde bereits 1994 ein Selbstbeobachtungsverfahren entwickelt, das sowohl die allgemeinen Standardkriterien der Forschungsevaluation als auch die besonderen Maßstäbe eines auf "problemorientierte Grundlagenforschung" festgelegten Forschungsprogramms berücksichtigen sollte. Ausgehend von der Prämisse, daß eine wirksame Evaluationspraxis auf einem gemeinsamen Grundverständnis der beteiligten Akteure beruht, ist im WZB in einem intensiven Diskussions- und Vermittlungsprozeß das Instrumentarium ("WZB-Erhebung") diskutiert, kritisiert und weiterentwickelt worden.

Es besteht aus einem qualitativen Teil, der sich an die Forschungseinheiten wendet. Dabei wurden in Form einer Selbstdarstellung insbesondere die Ziele des Forschungsprogramms, thematische Verknüpfungen mit anderen Abteilungen, externe Kooperationen sowie die Art und Weise der Vermittlung der Forschungsergebnisse

erfragt. Darüber hinaus wurden Einschätzungen eigener Stärken und Schwächen, die Beurteilung struktureller Rahmenbedingungen für die Forschung sowie künftige Orientierungen erkundet. Diese Informationen bildeten die Folie zur angemessenen Interpretation quantitativer Daten der Deskriptorenerhebung, die regelmäßig erhoben werden und sich im Zeitverlauf darstellen lassen. Auf Grundlage dieser Zeitreihen werden bestimmte Entwicklungstendenzen des Instituts sichtbar und interpretierbar. Die Deskriptoren sind zu unterscheiden nach Inputgrößen (z.B. Mittel, Stellen, Merkmale des wissenschaftlichen Personals), Throughputgrößen (z.B. Projektmerkmale, Kooperationen) und Outputgrößen (z.B. Publikationen, Drittmittelakquisitionen). Denn die Aussagekraft von Indikatoren hängt "von den Besonderheiten der Wissenschaftspraxis ab, in denen sich die Fachdisziplinen und Forschungsfelder sehr erheblich unterscheiden" (Neidhardt 1996).

Intendiert war mit diesem Verfahren, Aussagen über die institutionelle Leistungsfähigkeit, ihre Voraussetzungen und Bedingungen sowie die Weiterentwicklungsmöglichkeiten treffen zu können und einen umsetzungsorientierten Diskussionsprozeß einzuleiten. Von vornherein ausgeschlossen waren Erhebungen und Darstellungen individueller Leistungsdaten.

Selbst bei einem solchen Konzept von Selbstbeobachtungs- und Kontrollverfahren kann weder eine spezifische Innensicht und Problemwahrnehmung noch eine Ausblendung von Außenentwicklungen - gesellschaftliche und wissenschaftliche -, die die Forschung tangiert, vermieden werden. Insofern kann und sollte dieses Verfahren nicht externe Evaluationen ersetzen, sondern im Gegenteil in diesem Kontext verstanden und interpretiert werden.

Der Wissenschaftsrat hat in seinem Bewertungsbericht zum WZB sowohl die Aktivitäten des WZB-Beirats, der 1995 ein Audit aller Forschungseinheiten durchgeführt hat als auch die "Überlegungen zur Entwicklung von Leistungsindikatoren zur Dauerbeobachtung" positiv gewürdigt und empfohlen, mit anderen Blauen Liste-Einrichtungen und der WBL ein geeignetes methodisches Instrumentarium für die Selbstevaluierung der Blauen Liste zu erarbeiten (vgl. Wissenschaftsrat 1997, S. 58 f.).

III. Methodisches Vorgehen

Mit unserem Projekt soll ein Beitrag zur Diskussion über Ziele, Verfahren und Instrumente der Evaluation von Blaue Liste-Instituten geleistet werden. Weil uns insbesondere die Perspektive der Forschungsinstitute interessiert und wir versucht haben, ein möglichst breites Spektrum von Einschätzungen und Kommentaren zu erfassen, haben wir uns neben der Dokumentenanalyse auf leitfadenorientierte Interviews in ausgewählten Einrichtungen der WGL gestützt.

Die Interviews fanden in insgesamt 5 Forschungseinrichtungen statt - die Bibliotheken und Museen wurden aufgrund ihrer Aufgaben und ihrer Arbeitsweise, die sie erheblich von Forschungseinrichtungen unterscheiden, in dieser Studie nicht berücksichtigt. Zwar ist das Sample zu klein, um repräsentativ für die WGL zu sein - darauf kam es auch nicht an -, aber es gelang, bei der Auswahl das breite Spektrum der disziplinären Ausrichtung und unterschiedliche Ausprägungen des Forschungstypus - zwischen grundlagen- und anwendungsorientierter Forschung - zu berücksichtigen. Die Einrichtungen gehören zu den fünf verschiedenen Sektionen der WGL, und alle hatten zum Zeitpunkt der Interviews die Evaluation durch den Wissenschaftsrat erfolgreich absolviert.

Allen Einrichtungen, die sich an der Studie beteiligt haben, wurde eine Anonymisierung der Daten zugesagt. Im folgenden wird daher die Rede sein von *Institut A*, einer Einrichtung der Wirtschaftsforschung, und *Institut B*, einem pädagogischen Institut. *Institut D* erforscht und produziert technologische Artefakte, und sowohl *Institut C* als auch *Institut E* bewegen sich auf inter- bzw. multidisziplinär strukturierten Forschungsfeldern. Die beiden letztgenannten Institute sind naturwissenschaftlich orientiert, wobei *Institut E* zu einem geringen Anteil auch sozialwissenschaftliche Fragestellungen bearbeitet.

In jeder Einrichtung wurden insgesamt 12 Personen befragt, die entweder die verschiedenen, in einem Forschungsinstitut tätigen Personengruppen repräsentieren oder wichtige Funktionsträger sind. Im einzelnen handelte es sich jeweils um den Institutsdirektor, den Verwaltungsleiter, zwei AbteilungsleiterInnen, zwei wissenschaftliche MitarbeiterInnen, eine/n Doktorand/in, den Leiter oder die Leiterin einer wichtigen Serviceabteilung innerhalb der Forschungseinrichtung (beispielsweise das Labor oder die

Werkstätten), den Betriebs- oder Personalrat, die Gleichstellungsbeauftragte, den Sprecher der Mitwirkungsgremien (z.B. des "Wissenschaftlichen Rates") und um jemanden, der in der Einrichtung für die Forschungsplanung oder -organisation zuständig ist.

Der Leitfaden hat insgesamt drei Komplexe: Im ersten Teil der Interviews wurde nach zentralen Organisations- und Strukturmerkmalen der Einrichtungen gefragt, und im zweiten Teil wurden Einschätzungen zu Zielen und Verfahren von Evaluationen erkundet. Dabei kam es uns besonders darauf an, Vorschläge für einen zukünftigen Indikatorenkatalog zu eruieren. In einem dritten und letzten Teil haben wir unsere GesprächspartnerInnen gebeten, einzelne Aspekte der Evaluation durch den Wissenschaftsrates zu kommentieren, um dadurch eventuell notwendige Korrekturen der Beurteilungsmaßstäbe und -verfahren zu überprüfen.

Wie bereits erwähnt ging es uns nicht darum, repräsentative Meinungsbilder zu ermitteln. Unser Anliegen war es, in den ausführlichen Gesprächen mit den InterviewpartnerInnen differenzierte Kenntnisse über die spezifischen Strukturmerkmale des jeweiligen Instituts zu erhalten, die sich Außenstehenden nur schwer erschließen, aber in Evaluationen berücksichtigt werden müssen, sowie unterschiedliche Perspektiven und Erfahrungen unserer GesprächspartnerInnen kennenzulernen und mit ihnen verschiedene Gestaltungsmöglichkeiten von Evaluationsverfahren zu erkunden.

In dem folgenden Kapitel wird eine erste Zusammenfassung der Interviews vorgelegt. Wir konzentrieren uns auf fünf Themenkomplexe, die eine besondere Herausforderung für die Konzeption des zukünftigen Evaluationsverfahrens der WGL darstellen; nach dem Problemaufriß folgt eine (vorläufige) Darstellung der Hinweise und Anregungen, die wir in den Instituten erhalten haben. Wie in einem Zwischenbericht nicht anders möglich, gehen wir auf einzelne Aspekte und besondere Auffälligkeiten ein - eine differenzierte Analyse, die unter anderem auch kontroverse Positionen und Interessen der verschiedenen Personengruppen herausarbeitet, bleibt dem Endbericht des Projektes vorbehalten.

IV. Fragestellungen und Zwischenbefunde des Projekts

1. Aufgaben, Ziele und Organisation von WGL-Instituten

Die verschiedenen Einrichtungen der WGL sind, wie geschildert, ausgesprochen heterogen. Sieht man einmal von den Museen, Bibliotheken und denjenigen Einrichtungen ab, die als Serviceeinrichtungen klassifiziert werden, bleibt eine große Zahl von Forschungseinrichtungen, die bis auf die Finanzierung wenig gemeinsam haben. Die Einrichtungen weisen verschiedene Rechtsformen auf, es gibt eine große Spannbreite unterschiedlicher Disziplinen, und darüber hinaus lassen sich die einzelnen Institute keinem bestimmten Forschungstypus⁹ zuordnen: Weder betreiben sie überwiegend grundlagenorientierte Forschung wie die Max-Planck-Institute noch angewandte und "marktgängige" Forschung wie die Fraunhofer-Institute. Die Blaue Liste war von Anfang an eine "Art ökologischer Nische" (Hohn/Schimank 1990, S. 141), und auch bei vielen Instituten, die nach 1991 zur "Blauen Liste" hinzugekommen sind, handelt es sich um sehr heterogene Einrichtungen, die eine zunächst befristete Perspektive in diesem Sammelbecken gefunden haben.

Diese Heterogenität legt die Frage nahe, wie es bei so großen Unterschieden gelingen kann, ein Evaluationsverfahren und Instrumente zu entwickeln, die den vielen Varianten der Aufgaben- und Zielstellung von Blaue Liste-Instituten gerecht werden. Generell halten wir es für möglich, einen Rahmen und allgemeine Kriterien zu konstruieren, die dann allerdings spezifiziert werden müssen. Wir gehen davon aus, daß die angewendeten Indikatoren auf die jeweilige Einrichtung zugeschnitten bzw. dieser angemessen sein müssen: Beispielsweise wäre es ebenso falsch, ein eher anwendungsorientiertes Institut an der Zahl von Publikationen in referierten Journals zu messen wie es unangemessen wäre, von einem eher grundlagenorientierten Institut eine hohe Zahl von Industriekooperationen zu verlangen. Die Bewertungskriterien müssen, so unser Postulat,

⁹ Unter "Forschungstypus" verstehen wir mit Hohn/Schimank "... eine jeweils charakteristische Kombination sinnhafter Orientierungen und sozialstruktureller Einbettungen. So ist etwa die akademische Grundlagenforschung primär an Theoriefortschritten orientiert, und die primären Bezugsakteure solcher Grundlagenforscher sind die anderen Mitglieder der jeweiligen disziplinären oder subdisziplinären scientific communities. Im Unterschied dazu ist die Industrieforschung primär an technischen Innovationen von Produkten und Produktionsverfahren orientiert, und die primären Bezugsakteure der Forschungsabteilung eines Unternehmens sind andere Unternehmensabteilungen und die Kunden des betreffenden Unternehmens." (Hohn/Schimank 1990, S. 22)

der jeweiligen Aufgabe einer Forschungseinrichtung adäquat sein, um Verzerrungen zu vermeiden und um zu verhindern, daß die Institute an Maßstäben gemessen werden, die mit ihren Aufgaben und Zielen wenig zu tun haben.

Will man dieser Anforderung gerecht werden, so muß man in jedem Einzelfall das *Aufgabenprofil* eines Institutes möglichst genau erfassen. Wir unterscheiden zwischen drei Dimensionen des Aufgabenprofils, auf die im folgenden näher eingegangen werden soll: den Gegenstandsbereich, den Forschungstypus und den Anwendungskontext.

Zunächst ist es notwendig, die verschiedenen (programmspezifischen) Leistungen einer Einrichtung zu identifizieren, um dann, in einem zweiten Schritt, danach zu fragen, in welcher Weise sich diese Leistungen messen, bewerten oder vielleicht auch nur beschreiben lassen. Daher wurde auf der Grundlage von Dokumentenanalysen und von Interviews mit MitarbeiterInnen der genannten Forschungseinrichtungen exemplarisch versucht, einen kleinen Ausschnitt des Leistungsspektrums von WGL-Instituten zu ermitteln.

Zu Beginn wurden auf der Grundlage der Satzungen und zurückliegenden Jahresberichte der Einrichtungen sowie der Empfehlungen des Wissenschaftsrates die disziplinären Schwerpunkte, das Forschungsfeld und damit das Spektrum der zu bearbeitenden Forschungsfragen und -themen identifiziert. Auf diese Weise können erste Informationen über das jeweilige Aufgabenprofil des Institutes und jene "Besonderheiten der Wissenschaftspraxis" gewonnen werden, "in denen sich die Fachdisziplinen und Forschungsfelder (zum Beispiel im Hinblick auf Drittmittelbedarf, Publikationsregeln, Zitiergewohnheiten, Preishäufigkeiten etc.) sehr erheblich unterscheiden" (Neidhardt 1996). Da die interdisziplinäre Arbeitsweise als eine besondere Stärke der außeruniversitären Forschungseinrichtungen gilt, haben wir nicht nur die schriftlichen Quellen daraufhin untersucht, sondern auch in den Interviews die Formen der interdisziplinären Kooperation erfragt. Angesichts des breiten Aufgaben- und Themenspektrums der WGL-Institute gingen wir davon aus, auf beide Grundtypen der interdisziplinären Forschung zu stoßen: Sowohl auf die "interdisziplinäre Kooperation zur Lösung von Praxisproblemen" wie auf die "Kooperation mit dem Ziel interdisziplinärer Theoriebildung" (Blaschke 1986, S.168).

Darüber hinaus wurde in den Interviews das Aufgabenprofil der Einrichtungen unter folgenden Aspekten erörtert:

Wie erwähnt, bearbeiten die WGL-Institute ein Aufgabenspektrum zwischen Grundlagenforschung und anwendungsorientierter Forschung. 1995 wurde in einem Grundsatzpapier selbstbewußt festgehalten: "Während die Max-Planck-Gesellschaft sich überwiegend der zweckfreien reinen Grundlagenforschung verpflichtet fühlt und die Fraunhofer-Gesellschaft der angewandten, industrienahen Forschung dient, wobei sich beide als zentral organisierte, bundesweit operierende Gesellschaften präsentieren, sieht die Wissenschaftsgemeinschaft Bund-Länder-Institute ihre Stärke gerade in der Verbindung und ganzheitlichen Bearbeitung beider Aspekte und der Vermittlung und Verbreitung der Ergebnisse...".¹⁰

Die einzelnen Einrichtungen bearbeiten sowohl grundlagenorientierte als auch anwendungsorientierte Forschungsfragen. Man wird aber wohl davon ausgehen können, daß die WGL-Institute in der Regel keine "reine Grundlagenforschung" betreiben. Folgt man einer begrifflichen Unterscheidung von Jürgen Mittelstraß, bewegen sich die Forschungsaufgaben der WGL-Institute in einer Spanne zwischen "anwendungsorientierter Grundlagenforschung" und "produktorientierter Anwendungsforschung".¹¹

Die Ausgestaltung der beiden Schwerpunkte wird sich in jedem Institut unterscheiden. Offenkundig sind in den WGL-Instituten beide Forschungstypen mit unterschiedlichen Schwerpunktsetzungen vertreten, so daß häufig von Forschungsinstituten mit einer "überwiegend grundlagentheoretisch orientierten" oder einer "eher anwendungsorientierten" Ausrichtung gesprochen wird.

Eine wichtige Voraussetzung für erfolgreiche Evaluationen dürfte darin bestehen, zunächst das institutsspezifische Verhältnis von grundlagen- und anwendungsorientierten Forschungsarbeiten zu ermitteln, denn die unterschiedlichen Orientierungen haben

¹⁰ Der Text wurde von dem Grundsatzausschuß der Arbeitsgemeinschaft Blaue Liste erstellt, einer Vorläufereinrichtung der heutigen WGL (Grundsatzausschuss 1995, S. 24). Zum damaligen Zeitpunkt war die Namensfindung noch nicht abgeschlossen, die hier verwendete Bezeichnung "Wissenschaftsgemeinschaft Bund-Länder-Institute" konnte sich nicht durchsetzen.

¹¹ Reine Grundlagenforschung läßt nach Mittelstraß keine praktische Anwendung erwarten. Von den Ergebnissen anwendungsorientierter Grundlagenforschung werden langfristig Anwendungen erhofft, und produktorientierte Anwendungsforschung wird entweder im Hinblick auf Anwendungen oder in der Erwartung einer kurzfristigen Anwendung betrieben (vgl. Mittelstraß 1992).

Auswirkungen auf die Wahl der Bewertungskriterien bzw. die differenzierte Gestaltung der Output-Indikatoren: Beispielsweise kann es im Falle eines stark anwendungsorientierten Institutes sinnvoll sein, die Veröffentlichung der Forschungsergebnisse in breitenwirksamen oder adressatenorientierten Publikationen hoch zu bewerten, während die Leistungen einer eher grundlagenorientierten Einrichtung am besten durch die Zahl der Publikationen in referierten Zeitschriften dokumentiert werden. Denkbar ist jedoch ebenso, daß eine eindeutig dominierende Forschungsorientierung nicht erkennbar ist und demnach ein breites Spektrum von Forschungsleistungen bewertet werden muß.

Uns interessiert also das Verhältnis von grundlagen- und anwendungsorientierten Forschungsarbeiten in Instituten mit unterschiedlichen Aufgaben. Daher wurden fünf Institute ausgewählt, deren satzungsgemäßer Forschungsauftrag sich deutlich unterscheidet: Zwei Forschungsinstitute bearbeiten anwendungsorientierte Fragestellungen, zwei weitere Einrichtungen sollen sich laut Satzung mit grundlagenorientierten Forschungsfragen auseinandersetzen und bei einem Institut handelt es sich um eine Einrichtung, die als "Dienstleistungseinrichtung" für die Forschung bezeichnet wird. In diesen Einrichtungen haben wir die MitarbeiterInnen nach der Relevanz von Grundlagen- und anwendungsorientierter Forschung für die Erfüllung der Institutsziele befragt; einige Befunde werden im folgenden Abschnitt dargestellt.

Daneben untersuchen wir, welche externen Dienstleistungs- und Beratungsaufgaben die Institute wahrnehmen und in welcher Weise diese Leistungen im Rahmen von Evaluationen bewertet werden sollen bzw. können. Wir nehmen an, daß die Dienstleistungs- und Beratungsleistungen von erheblicher Bedeutung sind: Sie können (und sollten) zum einen, insbesondere in anwendungsorientierten Einrichtungen, eng mit den Forschungsarbeiten zusammenhängen und integraler Bestandteil des Forschungsprozesses sein. Die Dienstleistungen und Beratungen werden zum anderen aber auch aus wissenschaftspolitischen Gründen wichtig sein, denn es kann sich dabei um die von Bund und Ländern erwartete Übernahme von Aufgaben handeln, die von "überregionaler Bedeutung" und von "gesamtstaatlichem wissenschaftspolitischen Interesse" sind. Auch die Art der Beratungs- und Dienstleistungen wurde daher mit den MitarbeiterInnen ausführlich erörtert.

Mit obiger Feststellung wird bereits angedeutet, daß das Aufgabenprofil eines WGL-Instituts sowohl durch die kognitive Struktur des Forschungsfeldes als auch durch

forschungsexterne Einflüsse bestimmt wird. Daher haben wir die MitarbeiterInnen der Forschungsinstitute nach den Beziehungen zu Politik, Wirtschaft und Gesellschaft befragt. Besonders wichtig sind - ähnlich wie bei anderen außeruniversitären, staatlich finanzierten Einrichtungen - die Beziehungen zu den Universitäten: Einerseits wird von den WGL-Instituten erwartet, daß sich ihre Forschungsaufgaben deutlich von universitären Aufgaben abgrenzen und sie in "Forschungsnischen" tätig sind, die von den Hochschulen nicht oder nicht vergleichbar erfolgreich bearbeitet werden können. Andererseits wird von den WGL-Einrichtungen eine enge Kooperation mit den Hochschulen erwartet. Insofern hängt die Leistungsfähigkeit eines WGL-Institutes, so unsere Annahme, auch mit erfolgreichen Abstimmungsprozessen und Kooperationsstrukturen zusammen. Auf die Erfahrungen und Einschätzungen unserer GesprächspartnerInnen wird im folgenden Abschnitt eingegangen.

1.1 Aufgaben und Ziele: Entwicklung und Behauptung eines eigenen Profils

Die WGL-Institute bearbeiten, wie geschildert, sowohl grundlagen- als auch anwendungsorientierte Forschungsfragen. Die Verbindung der beiden Forschungstypen stellt eine besonders schwierige Aufgabe dar. Ein Problem besteht darin, Organisations- und Steuerungsformen zu finden, die nicht nur ein bloßes Nebeneinander von Grundlagenforschung und angewandter Forschung, sondern eine funktionale Integration ermöglichen und unterstützen. Wie jüngere Studien gezeigt haben, handelt es sich bei einer "Kopräsenz" der beiden Forschungsorientierungen um eine eher instabile Konstruktion: Sie führt tendenziell zu einer Konkurrenz zwischen grundlagen- und anwendungsorientierten Forschern und zu einem Verdrängungsdruck, bei dem häufig die - nicht durch Industriepartner unterstützte und zusätzlich finanzierte - Grundlagenforschung unterliegt (vgl. Hohn/Schimank 1990: 69 ff., Gläser/Meske 1996).

Wir haben von den MitarbeiterInnen der von uns besuchten Forschungseinrichtungen erfahren wollen, ob und in welcher Weise die Verbindung von Grundlagen- und anwendungsorientierter Forschung das Profil des jeweiligen Institutes prägt und wie diese Verbindung aktiv gestaltet wird. Wir gehen davon aus, daß eine bewußte und offensive Ausgestaltung des Verhältnisses der beiden Forschungsorientierungen in den Instituten ein Hinweis auf entwickelte Selbststeuerungsfähigkeiten ist. Darüber hinaus ist, wie oben ausgeführt, die Klärung der Gewichtung bzw. Schwerpunktsetzung von Grund-

lagen- und anwendungsorientierter Forschung eine wichtige Voraussetzung für die Ermittlung eines adäquaten Indikatorensatzes.

In den Gesprächen haben sich auffällige Unterschiede hinsichtlich der Integration von grundlagen- und anwendungsorientierter Forschung gezeigt. Bemerkenswert ist zunächst einmal, daß in beiden eher anwendungsorientierten Instituten der Stellenwert der grundlagenorientierten Forschungen hervorgehoben wurde. Im *Institut A*, das sich auf empirische Wirtschaftsforschung konzentriert, wiesen mehrere MitarbeiterInnen darauf hin, daß Datenbanken gepflegt werden, die eine wichtige Basis für überwiegend grundlagenorientierte Forschungsarbeiten darstellen. Darüber hinaus wurde kommentiert, daß die Grundlagenforschung zwar vom Volumen her eher gering, aber dennoch wichtig für die Generierung neuer Forschungsfragen sei. Im *Institut B*, das Forschungen an der Schnittstelle von Fachdidaktik, Pädagogik und Psychologie betreibt, wurde der quantitative Anteil von grundlagenorientierten Forschungen sogar auf 40-50% geschätzt, und mehrere Interviewpartner bezeichneten grundlagen- und anwendungsorientierte Forschung als die "zwei Säulen" des Instituts.

Überraschend war, daß auch die MitarbeiterInnen des *Institut E* - einer Einrichtung, die von der WGL als "Serviceeinrichtung" bezeichnet wird - den grundlagenorientierten Arbeiten große Bedeutung beimaßen. Es gebe, wurde uns erläutert, eine enge Wechselwirkung zwischen der Züchtung von Kristallen und den Forschungsarbeiten: Um die Qualität der gezüchteten Kristalle zu beurteilen, sind Forschungsarbeiten notwendig, die wiederum eine Grundlage für die Verbesserung der Kristalle darstellen oder dabei helfen sollen, in einem anschließenden Züchtungsverfahren die gewünschten Eigenschaften herzustellen. Offensichtlich stellt das *Institut E* ein gutes Beispiel für einen neuen Typus der Wissensproduktion dar: Forschungserkenntnisse können unmittelbar in neue "Produkte", d.h. in die Züchtung von Kristallen umgesetzt werden, während die mit großem wissenschaftlich-technischen Know-how hergestellten Kristalle ihrerseits auf die Grundlagenforschung zurückwirken und neue Forschungsfragen generieren.¹²

¹² Ähnlich charakterisieren Gibbons et al. den neuen Typ der Wissensproduktion: "It is characterised by a constant flow back and forth between the fundamental and the applied, between the theoretical and the practical. Typically, discovery occurs in contexts where knowledge is developed for and put to use, while results - which would have been traditionally characterised as applied - fuel further theoretical advances." (Gibbons et al. 1994, S. 19)

Im Unterschied zu diesen drei Einrichtungen fielen die Urteile über das Zusammenwirken von Grundlagen- und Anwendungsforschung in den beiden eher grundlagenorientierten Forschungseinrichtungen, die ebenfalls anwendungsorientierte Forschungsfragen bearbeiten, ambivalent aus. Zwar wurde die Koexistenz von grundlagen- und anwendungsorientierter Forschung begrüßt, aber mehrere GesprächspartnerInnen kritisierten, daß die beiden Forschungsrichtungen weniger integriert als vielmehr parallel und unverbunden bearbeitet werden. In beiden Einrichtungen gibt es Meinungsunterschiede über die Frage, welche Relevanz Grundlagen- und Anwendungsforschung in dem jeweiligen Forschungsfeld haben und in welcher Weise sich die Einrichtung hinsichtlich der Umsetzung der Forschungsergebnisse und der Kooperation mit externen Akteuren engagieren sollte.

Es verbietet sich jedoch - nicht nur wegen der kleinen Zahl von zwei Einrichtungen - daraus zu schließen, daß eher grundlagenorientierte Institute größere Probleme mit der Integration von anwendungsorientierten Forschungsfragen haben als umgekehrt. Vielmehr haben die Interviews gezeigt, daß die Spannungen heterogene Ursachen haben:

- In beiden Einrichtungen wird zumindest ein Teil der Grundlagenforschung in eigenen Forschungseinheiten (z.B. Abteilungen) organisiert. Die damit verbundene unterschiedliche Anbindung der Abteilungsleiter an die Universitäten sowie die Orientierung an verschiedenen Adressaten der Forschungsleistungen, an unterschiedlichen Reputationskriterien und Karrieremustern verstärken offenkundig Tendenzen zur Separierung;
- beide Einrichtungen sind 1992 gegründet worden, sie sind also relativ jung und sind als jeweils einzige auf inter- bzw. multidisziplinär strukturierten Forschungsfeldern tätig. Beide Einrichtungen entwickeln sich noch, und insofern können die erwähnten Spannungen zwischen den verschiedenen Orientierungen auch ein Nebeneffekt der Profilierungsbemühungen der beiden WGL-Institute im Kontext unterschiedlicher Fachdisziplinen, Wissenschaftseinrichtungen und internationaler Forschungsaktivitäten sein;
- darüber hinaus bezogen sich mehrere MitarbeiterInnen der beiden Einrichtungen explizit auf das Forschungsverständnis der Vorgängereinrichtungen, d.h. auf die ehemaligen Institute der Akademie der Wissenschaften der DDR. Forschung sollte

sich ihrer Meinung nach, gerade in den Forschungsfeldern ihrer Einrichtungen, auf wichtige Problemlagen und gesellschaftlich relevante Forschungsfragen konzentrieren. Ihre Position deckt sich durchaus mit einem Forschungsbegriff, den auch MitarbeiterInnen anderer Institute vertreten, und mit den Erwartungen seitens des Bundes, der Länder und des Wissenschaftsrates an diese Einrichtungen. Dennoch scheinen in beiden Fällen die bekannten Ost-West-Konflikte das Verhältnis der unterschiedlichen Forschungsorientierungen spannungsreich zu gestalten: Hier entstand eine unglückliche Konstellation von "alten" MitarbeiterInnen, die eher anwendungsorientierte Ziele vertreten, und "neuen" WissenschaftlerInnen, die mehr an grundlagenorientierten Forschungsfragen interessiert sind.

Auffallend ist, daß viele GesprächspartnerInnen einen recht weit gefaßten Begriff von "anwendungsorientierter Forschung" haben. Wir haben darunter zunächst die Orientierung an wissenschaftsexternen Problemen und die Nutzung der Erkenntnisse durch außerwissenschaftliche Interessenten verstanden.¹³ Für unsere GesprächspartnerInnen wird anwendungsorientierte Forschung darüber hinaus durch den Entstehungszusammenhang definiert: Anwendungsorientierte Forschungen würden häufig durch einen externen Akteur angeregt oder von diesem finanziert. Sie verstanden ferner auch Vermittlungs- und Umsetzungsprozesse als Bestandteile von "anwendungsorientierten Forschungen". Forschungs- und Beratungsleistungen gehen also fließend ineinander über, was - jedenfalls zum Teil - den großen Stellenwert von Beratungs- und Dienstleistungsaufgaben in den WGL-Instituten erklärt. Beispielfhaft soll ein kleiner Ausschnitt dargestellt werden:

- Ein Institut reagiert auf Anfragen aus der Bevölkerung und unterhält ein kleines Beratungszentrum;
- eine Forschungseinrichtung stellt als Ergebnis der Forschungsarbeiten Kristalle her, die in enger Abstimmung mit den externen "Abnehmern", also Industrieunternehmen, Forschungseinrichtungen und Hochschulen, produziert, geprüft und verbessert werden;

¹³ Zur Diskussion der Definitions- und Abgrenzungsprobleme von "Grundlagenforschung" und "Anwendungsforschung" vgl. Gläser/Meske (1996, S. 36 ff.).

- eine Einrichtung betreibt Politikberatung und arbeitet zur Umsetzung der Forschungsergebnisse eng mit Forst- und Landwirten zusammen;
- ein Institut entwickelt didaktische Konzepte für den mathematisch-naturwissenschaftlichen Unterricht, führt Lehrerfortbildungen durch und organisiert verschiedene Schülerwettbewerbe, und
- eine Einrichtung sieht eine wesentliche Aufgabe in der Politikberatung, die auf empirisch gewonnenen Forschungserkenntnissen basiert, und unterstützt Verwaltung, Wissenschaft und Wirtschaft durch regelmäßige Berichte und Gutachten. Außerdem sammelt und pflegt diese Einrichtung Datensätze, die anderen Wissenschaftsinstitutionen zur Verfügung gestellt werden.

Es ist ein nicht unbeträchtlicher Zeit- und Organisationsaufwand dazu erforderlich, diese Leistungen zu erbringen. Zum Teil werden einzelne Personen oder Abteilungen damit beauftragt, und teilweise gehören Information, Beratung und Unterstützung zum Tätigkeitsprofil mehrerer MitarbeiterInnen, für die spezielle Kompetenzen ausgebildet werden oder vorhanden sein müssen. Diese Leistungen sind also nicht etwa "zusätzlich" und wissenschaftsfremd, sondern ein wichtiger Bestandteil der Forschungsarbeiten. Sie werden zum Teil durch externe Akteure angeregt, von ihnen gefördert oder finanziert, und es ist durchaus wichtig, ob die Zuwendungsgeber diese Aktivitäten unterstützen oder sogar - etwa aus regional- und strukturpolitischen Gründen - fordern.

Als ein weiterer Aufgabenbereich wurde die Gestaltung der Beziehungen zu den externen Akteuren genannt, unter denen die Universitäten eine besonders wichtige Rolle spielen. Die WGL-Institute unterscheiden sich von den Universitäten in der Regel nicht nur durch eine ausgeprägtere Anwendungsorientierung. Im Unterschied zu den Universitäten ist in den meisten Instituten ein großer Teil des Forschungspersonals kontinuierlich tätig; die MitarbeiterInnen können langjährige Forschungserfahrungen, theoretisches Wissen und technisches Know-how vorweisen, was zugleich eine wichtige Voraussetzung für mehrjährig angelegte Forschungsprojekte darstellt (aber bei mangelnder Mobilität auch Probleme mit sich bringen kann); und die Institute haben Organisations- und Entscheidungsstrukturen, die ausschließlich auf die Erfordernisse der Forschung zugeschnitten sind.

Viele leitende Wissenschaftler haben häufig zugleich eine Professur an einer benachbarten Universität inne. Dadurch wird eine Voraussetzung für die Vertiefung und Ergänzung des fachlichen Profils geschaffen, und zugleich kann die Personalunion ein Problem lösen, das sich für viele außeruniversitäre Forschungseinrichtungen stellt: Eine erfolgsversprechende Rekrutierung des wissenschaftlichen Nachwuchses ist ohne direkten Zugang zu den Universitäten kaum möglich.

Die meisten Institute haben einen Kooperationsvertrag mit einer Universität - im Berliner Raum auch mit mehreren Universitäten - abgeschlossen, in dem beispielsweise die Modalitäten von gemeinsamen Berufungen oder Forschungskooperationen, aber auch Lehrverpflichtungen festgehalten werden. Die Universitäten haben ihrerseits in mehreren Fällen die Möglichkeit, als Mitglieder der Aufsichtsgremien (Vorstand oder Kuratorium) auf die Entwicklung von WGL-Instituten Einfluß zu nehmen, und nicht zuletzt waren an den Gründungskomitees, die Empfehlungen für die Neugründung von Instituten abgaben, universitäre Gutachter beteiligt.

Trotz dieser formal engen Bindungen haben unsere GesprächsteilnehmerInnen die Qualität und Intensität der Forschungskooperationen nicht einheitlich bewertet: Neben ausgesprochen positiven Urteilen gibt es auch selbstkritische Einschätzungen, daß die Kooperation ausbaufähig sei. Insbesondere die Institute in den neuen Bundesländern befinden sich offensichtlich noch in einer Phase, in der universitäre Ansprechpartner identifiziert und Formen der Zusammenarbeit aufgebaut werden müssen. In einem Institut wurden die geringen Kooperationen mit der "Zurückhaltung" der Universität begründet - eine Einschätzung, die die Gutachtergruppe des Wissenschaftsrates übrigens teilte.

Alles in allem gibt es aus der Sicht der meisten GesprächspartnerInnen eine funktionierende Arbeitsteilung zwischen Universitäten und Blaue Liste-Instituten, die darauf beruhe, daß die WGL-Institute subsidiär auf Gebieten tätig seien, deren Bearbeitung für die Hochschulen aus inhaltlichen wie strukturellen Gründen nicht oder nur schwer möglich wäre. Diese selbstbewußte Einschätzung findet offenkundig die Zustimmung des Wissenschaftsrates, der zur "Halbzeit" der Bewertungen der Blaue Liste-Institute nur in einem Fall die Eingliederung in eine Universität empfahl. Als Argumente für eine organisatorische Selbständigkeit werden neben einer "überzeugenden Tradition" auch die "Mittelpunkts- und Mittlerfunktion", die "Singularität des Forschungs-

gebietes", die Langfristigkeit der Forschungsaufgaben und die erforderliche interdisziplinäre Zusammenarbeit genannt.¹⁴

Zusammenfassend und im Hinblick auf ihre Ziele und Aufgaben sei festgehalten, daß die Institute der WGL mit Erwartungen und Anforderungen von sehr unterschiedlichen Akteuren konfrontiert werden, von denen keiner vernachlässigt werden darf. So stehen die Leistungserwartungen der scientific community neben den Forderungen nach politischer Beratung, und die Wünsche der Anwender und Praxisvertreter müssen ebenso berücksichtigt werden wie die Kriterien des Wissenschaftsrates. Die Herausbildung und Stabilisierung ihrer Identität gestaltet sich schwierig: Im günstigen Fall halten die MitarbeiterInnen diese breite Spanne von Aktivitäten und Forderungen für ein positives und vielversprechendes Merkmal des jeweiligen Institutsprofils, in anderen Instituten dagegen dominieren die Befürchtungen, daß es sich bei der geschilderten Konstellation um einen "Spagat" handelt, der zu einer kaum zu bewältigenden Zerreißprobe werden könnte. Alle Einrichtungen werden sich auch zukünftig mit der Gestaltung dieser Spannungen offensiv auseinandersetzen müssen; dabei kann die kontinuierliche Beratung und Unterstützung durch die Beiräte eine wichtige Hilfe sein.

2. Organisation und Steuerung von WGL-Instituten

Über die Untersuchung des *Aufgabenprofils* hinaus gibt es mehrere Gründe, auch die *Organisationsstruktur* der WGL-Institute besonders aufmerksam zu betrachten.

Wie Renate Mayntz deutlich gemacht hat, ist die Aufbau- und Ablauforganisation von großer Bedeutung für die Leistungsfähigkeit einer Forschungsorganisation. Mayntz stimmt dem kontingenztheoretischen Ansatz der Organisationsforschung zu, der davon ausgeht, daß zwischen der Struktur einer Organisation und ihrem Aufgabentypus ein Entsprechungsverhältnis besteht. Es gibt also "angemessene" oder "unangemessene" Organisationsformen, und sie hält fest, daß "... die leistungsfähige Organisationsform

¹⁴ Vgl. dazu die "Halbzeitbilanz" von Winfried Benz, dem Generalsekretär des Wissenschaftsrates, in der Deutschen Universitätszeitung (Benz 1998).

der Aufgabenstellung und Größe einer Einrichtung, den Besonderheiten ihres Personals und ihrer apparativen Ausstattung angemessen sein muß".¹⁵

Allerdings ist keine Organisationsform unproblematisch, so ihre Kernthese, denn jede organisatorische Lösung hat es mit Grundspannungen zu tun, deren Bewältigung eine Daueraufgabe ist. Die Grundspannungen charakterisiert sie als Interessendivergenzen und Ambivalenzen, als Spannungen zwischen betriebsförmiger Organisation und schöpferischer Forschungstätigkeit. Aus der Steuerungsperspektive haben die Leiter der Forschungseinrichtungen die Aufgabe, "... jenen Schnittpunkt zu finden, der das für die individuelle Leistung förderliche Maß an Autonomie mit dem für die kollektive Produktivität nötigen Maß an Organisation kombiniert" (Mayntz 1985, S. 37). Es gibt jedoch keine "richtigen" Organisations- und Personalentscheidungen, mit denen die Konflikte und Spannungen ein für alle Mal gelöst werden, da die Grundspannungen unaufhebbar sind - jede Problemlösung ist zugleich die Quelle neuer Probleme. Außerdem können auch "angemessene" Organisationsentscheidungen die Leistungsfähigkeit einer Forschungsorganisation nicht langfristig sicherstellen, denn die zentralen Spannungen müssen fortlaufend bewältigt werden, da jede Einrichtung ein komplexes und dynamisches Gebilde in einer sich verändernden Umwelt ist.

Aus diesen Überlegungen ergeben sich wichtige Folgerungen für die Gestaltung von Evaluationsverfahren. Angesichts der hohen Bedeutung von Organisationsentscheidungen empfiehlt sich eine Überprüfung der jeweiligen, in den Instituten gefundenen strukturellen Lösungen, mit der eine Korrektur der getroffenen Entscheidungen veranlaßt werden kann. Diese Überprüfung sollte ein regelmäßig und kontinuierlich praktizierter Bestandteil von Evaluationen sein, um zu sichern, daß auf Veränderungen innerhalb der Organisation und auf Umweltveränderungen angemessen reagiert wird. Naheliegend ist, daß nicht nur die Organisationsstruktur der Institute, sondern auch die Managementfähigkeiten der Institutsleiter beurteilt werden, deren Aufgabe die kontinuierliche Gestaltung und Bewältigung der strukturellen Grundspannungen ist - die Organisationsentscheidungen und der Führungsstil der Institutsleiter sind daher von erheblicher Bedeutung für die Leistungsfähigkeit "ihrer" Institute.

¹⁵ Mayntz (1985, S. 32). In ihrer grundlegenden Studie werden erstmals Probleme der Forschungsorganisation in einem spezifischen Forschungssektor untersucht: den außeruniversitären, öffentlich finanzierten Forschungseinrichtungen. In jüngerer Zeit erschienen zu dieser Problematik zum Teil bereits genannte Untersuchungen von Hohn/Schimank (1990), Gläser/Meske (1996), Hasse (1996), Röbbcke (1997) und Hohn (1998).

Man wird allerdings in Zukunft eine Überprüfung der Organisationsstrukturen nicht nur *externen Evaluationen* überlassen dürfen, die in großen Abständen stattfinden und unter hohem Zeitdruck stehen. Externe Evaluationen bieten eine Möglichkeit, die Selbststeuerungsfähigkeiten einer Forschungseinrichtung zu kontrollieren, also festzustellen, ob und in welcher Weise sich ein Institut mit den Organisationsstrukturen auseinandersetzt, welche Spannungen und Probleme identifiziert und welche Weiterentwicklungsoptionen nicht nur diskutiert, sondern auch umgesetzt werden. Mit einer ausführlichen Bewertung einzelner Aspekte kann eine externe Evaluation schnell überfordert sein.

Interne Evaluationen dürften dagegen zur Feinbewertung der Organisationsstrukturen besser geeignet sein, wenn die Beiräte der Einrichtungen beteiligt sind. Die Beiräte haben den Vorzug, daß sie (in der Regel) mit der institutionellen Geschichte und den strukturellen Besonderheiten der Einrichtung vertraut sind. Insofern bringen die Beiräte gute Voraussetzungen mit, um begründete Urteile über die Angemessenheit von Strukturentscheidungen abgeben zu können. Die Vertrautheit mit den Institutsstrukturen wirkt sich vor allem dann als Vorteil aus, wenn institutionelle Vorkehrungen getroffen werden, die zusätzlich gewährleisten, daß innovative Potentiale bei Beratung und Bewertung zum Tragen kommen, z.B. durch befristete Amtszeiten von Beiratsmitgliedern.

Diese Vertrautheit ist nicht zuletzt deshalb wichtig, weil die Bewertungen und Empfehlungen - je nachdem, um welche Organisations- und Strukturmerkmale es sich handelt - gravierende Folgen haben können. Wenn ein Beirat die nachlässige *Aus- und Weiterbildungspraxis* einer Einrichtung kritisiert, wird diese Kritik für das Institut zwar schmerzhaft, aber verkraftbar sein. Dagegen wird es folgenreicher sein, wenn als Ergebnis einer Evaluation festgehalten wird, daß die *Abteilungsgliederung* den derzeitigen Forschungsaufgaben nicht (mehr) angemessen ist. Eine derartige Feststellung bzw. die Empfehlung zur Umstrukturierung muß das Ergebnis einer diffizilen Erörterung sein, in deren Rahmen erwogen wird, ob es sich bei den diskutierten Problemen um (möglicherweise temporäre) Begleiterscheinungen einer durchaus "angemessenen" Organisationsform handelt oder ob eine Umstrukturierung sinnvoll ist, die andere Probleme und neue Spannungen mit sich bringen wird.

Offen ist, *welche Merkmale* der Organisationsstruktur in Evaluationen bewertet und nach *welchen Kriterien* die Strukturmerkmale beurteilt werden sollen. Die Forschungseinrichtungen müssen organisatorische Lösungen für - jeweils unterschiedlich ausgeprägte - strukturelle Spannungen entwickeln: Beispielsweise müssen die Blaue-Liste-Institute grundlagen- und anwendungsorientierte Forschungsaufgaben verbinden; sie sollen inter- bzw. multidisziplinäre Forschungsfragen bearbeiten; ihre Forschungsfragen stehen im Spannungsfeld von Anforderungen der scientific community, des Staates und der Nutzer; und weil es sich um betriebsförmige Organisationen handelt, gibt es eine grundsätzliche Spannung zwischen den Organisationszielen und den individuellen Bedürfnissen und Interessen der MitarbeiterInnen. Eine Aufgabe von Evaluationen ist es, die Leistungsfähigkeit eines Institutes dahingehend zu überprüfen, welche organisatorischen Lösungen in der Einrichtung gefunden, erfolgreich praktiziert und weiterentwickelt werden.

Im Vorfeld der Evaluation einer Forschungseinrichtung muß erwogen werden, welche Organisations- und Strukturmerkmale begutachtet werden sollen. Nicht alle Aspekte lassen sich aufgrund theoretischer Überlegungen erschließen: Es wird drängende Probleme geben, die jeweils mit den Besonderheiten eines Instituts zusammenhängen, und daneben wird es Aufgabenbereiche geben, bei denen kontrovers diskutiert wird, ob sie Gegenstand von Evaluationen sein sollten - dazu gehört beispielsweise die Frauenförderung. Wir haben daher in den Gesprächen mit den Leitern und den MitarbeiterInnen der Forschungseinrichtungen ausführlich über Organisationsfragen gesprochen und zu ermitteln versucht, welche Gestaltungsmerkmale aus ihrer Sicht für die Leistungsfähigkeit ihrer Einrichtung relevant sind und Gegenstand von Evaluationen sein sollten.

Ebenfalls haben wir in den Gesprächen die oben erwähnte zweite Frage zu klären versucht, nach welchen Kriterien die einzelnen Merkmale beurteilt werden sollen. Derzeit basieren die Bewertungen überwiegend auf Erfahrungswissen und gemeinsam geteilten Einschätzungen und Überzeugungen der Gutachter, denn bisher haben sowohl die Wissenschafts- als auch die Organisationsforschung wenig dazu beigetragen, tragfähige Beurteilungsmaßstäbe zu entwickeln - selbst der Zusammenhang zwischen den Forschungsaufgaben und der "angemessenen" Organisationsform einer Forschungseinrichtung ist erst ansatzweise geklärt.

In den Interviews haben wir eine Reihe von Merkmalen erörtert, die auch im Rahmen der aktuellen WGL-Evaluation durch den Wissenschaftsrat beurteilt werden, unter anderem gehören die Gestaltung abteilungsübergreifender Kooperationen und die Befristungspraxis dazu. Andere Aspekte, wie die Frage nach Verfahren der zentralen Ressourcenallokation, gingen darüber hinaus. Von diesen Gesprächen versprochen wir uns Anregungen bei der Identifizierung von wesentlichen Organisationsmerkmalen und -problemen, aber auch Aufschluß über die Akzeptanz des derzeitigen Evaluationsverfahrens und Hinweise darauf, in welcher Weise die bestehenden Instrumente ergänzt oder modifiziert werden sollten. Keinesfalls ging es um die Ermittlung von Urteilen über Institute, sondern um die Entwicklung von Beurteilungsmaßstäben und -verfahren.

2.1 Organisation, Planung und Steuerung als Gegenstand der Evaluation

Wir haben in den Instituten, wie gesagt, eine ganze Reihe von Organisationsmerkmalen erörtert. An dieser Stelle kann darauf nicht umfassend eingegangen werden, daher konzentrieren wir uns auf eine kleine Auswahl.

1. Die Organisation von Abteilungen und abteilungsübergreifenden Kooperationen. Ein Problem, für das alle Forschungsinstitute eine Lösung finden müssen, ist die Frage der Binnenorganisation. Abhängig von der *Größe* des Institutes kann es sich beispielsweise in Forschungsschwerpunkte oder Abteilungen mit Projekten als kleinsten Forschungseinheiten gliedern. Eine weitere Variable ist die *Aufgabe*: Die Organisationsstrukturen müssen beispielsweise die Integration von Grundlagen- und anwendungsorientierten Forschungen oder die interdisziplinäre Bearbeitung von Forschungsfragen ermöglichen, wenn nicht sogar fördern.

Eine wichtige Entscheidung über die Organisationsgestaltung betrifft die Kriterien, nach denen innerhalb der Einrichtungen eine Differenzierung der Abteilungen vorgenommen wird.¹⁶ Grundsätzlich ist es möglich, Abteilungen nach verschiedenen Kriterien, zum Beispiel nach disziplinären Gesichtspunkten, Methoden, Theorien,

¹⁶ Alle fünf Institute sind nach dem Abteilungsmodell mit einer Direktorialverfassung strukturiert. Im Falle einer Einrichtung werden die Untereinheiten zwar als "Institut" bezeichnet, aber da die Forschungen in ein gemeinsames Konzept integriert sind, handelt es sich auch hierbei um ein Abteilungsmodell. Nachfolgend wird daher generell von "Instituten" und "Abteilungen" gesprochen.

"Produkten" (vor allem im naturwissenschaftlich-technischen Bereich) oder nach Teilfragen zu gliedern. Dabei wird man versuchen, sich möglichst an dem Aufgabentypus der Forschungseinrichtung zu orientieren.

Mayntz unterscheidet drei Aufgabentypen. Es gibt Institute mit einer *segmentären Aufgabenstruktur*, d.h. mit Forschungsaufgaben, die sich in Teilfragen zerlegen lassen, die parallel und unverbunden bearbeitet werden können. Bei einer *sequentiellen Aufgabenstruktur* werden einzelne Teile eines Forschungsprogrammes nacheinander bearbeitet, und bei der häufig vorkommenden "*pooled interdependence*" lassen sich Forschungsfragen zwar in Teilfragen zerlegen, diese müssen aber wieder vernetzt werden. Institute mit einer segmentären Aufgabenstruktur lassen sich demnach so gliedern, daß zwischen den Abteilungen keine sachlichen Abhängigkeiten bestehen; die anderen beiden Strukturtypen sind durch Abhängigkeitsmuster zwischen den Abteilungen gekennzeichnet und erfordern daher einen Koordinations- und Kooperationsaufwand. Wie hoch die zu erbringenden Steuerungsleistungen sind, hängt damit zusammen, nach welchen Prinzipien die Abteilungen gegliedert sind: "Die Notwendigkeiten abteilungsübergreifender Kooperation ... hängen von der Art der sachlichen Abhängigkeitsbeziehungen zwischen der Arbeit verschiedener Abteilungen und damit einerseits von der oder den Forschungsaufgaben des Instituts und andererseits von den Kriterien ab, die der Abteilungsgliederung zugrunde liegen." (Mayntz 1985, S. 67)

Dieser nicht ganz einfache Zusammenhang kann am Beispiel der von uns besuchten Institute verdeutlicht werden. Ein Institut ist durch einen segmentären Aufgabentypus gekennzeichnet, der durch eine stark gegenstandsbezogene Binnengliederung abgebildet wird. Die einzelnen Abteilungen beschäftigen sich mit verschiedenen "Gegenständen" oder Themen, sie sind interdisziplinär strukturiert und eine Differenzierung zwischen eher grundlagen- und eher anwendungsorientierten Fragestellungen ist nicht erkennbar. Diese Gliederung hat sich nach Darstellung der MitarbeiterInnen im Prinzip bewährt, und eine hierarchische Koordination abteilungsübergreifender Arbeiten scheint kaum notwendig zu sein. Ergänzend sei angemerkt, daß der reibungslose Aufbau abteilungsübergreifender Projektteams offensichtlich auch mit entsprechenden Außenanforderungen zusammenhängt, da es sich um Drittmittelprojekte handelt, deren Akquisition und erfolgreiche Bearbeitung eine hohe Bedeutung haben. Gleichwohl muß auch diese Einrichtung immer wieder Anpassungen an die sich in rasantem

Tempo wandelnden Forschungsaufgaben unternehmen und überprüfen, ob der Abteilungszuschnitt dem aktuellen Forschungsspektrum angemessen ist.

Ein anderes Institut bearbeitet interdisziplinäre Forschungsfragen und ist durch den Aufgabentypus der "pooled interdependence" gekennzeichnet. Die Abteilungsgliederung ist nach disziplinärem Gesichtspunkten vorgenommen worden. Die Abteilungen stehen im Mittelpunkt des wissenschaftlichen Lebens, und sowohl die Rekrutierung der wissenschaftlichen MitarbeiterInnen und des wissenschaftlichen Nachwuchses als auch der Reputationsgewinn innerhalb des Forschungsfeldes sind disziplinär organisiert. Daher ist die Organisation von Interdisziplinarität eine ständige Herausforderung: Für die Etablierung formaler, abteilungsübergreifender Kooperationsstrukturen und für die Entwicklung eines tragfähigen Forschungsprogramms sind hohe Steuerungsleistungen notwendig. Nach Ansicht unserer GesprächspartnerInnen - und der Gutachtergruppe des Wissenschaftsrates - war das Institut dabei bisher erfolgreich, und die entwickelten Kooperationsformen haben sich bewährt.

Am Beispiel eines anderen Institutes wird deutlich, daß das Forschungsspektrum die Abteilungsgliederung keineswegs determiniert. In dieser Einrichtung, die ebenfalls dem Aufgabentypus der "pooled interdependence" zuzuordnen ist, gibt es Abteilungen, die zum Teil thematisch und zum Teil disziplinär organisiert sind. Das führt dazu, daß es eher anwendungsorientierte Abteilungen neben stark universitär bzw. grundlagenorientierten Abteilungen gibt. Die Konstruktion ist offensichtlich ausgesprochen spannungsreich, und sie wurde von mehreren MitarbeiterInnen für eine unbefriedigende organisatorische Umsetzung des Forschungsfeldes gehalten. Die unterschiedlichen Orientierungen der Abteilungen erschweren eine Entwicklung abteilungsübergreifender Kooperationsstrukturen, und auch der Versuch der Leitung, durch Anreizstrukturen - etwa zusätzliche finanzielle und personelle Ressourcen - abteilungsübergreifende Projekte zu fördern, ist in den Abteilungen bisher auf wenig Resonanz gestoßen.

Abteilungsübergreifende Kooperationen haben also eine hohe Bedeutung. Sie sind in Instituten mit multidimensionalen Abhängigkeitsbeziehungen zwischen den Abteilungen sachlich notwendig und können darüber hinaus auch in Instituten mit einer segmentären Aufgabenstruktur zu erwünschten Synergieeffekten führen. Vermutlich will der Wissenschaftsrat im Rahmen der derzeitigen WGL-Evaluierung aus diesen Gründen wissen, wie ein Institut gegliedert ist, ob es eine abteilungsübergreifende

Forschungsplanung und -projekte sowie gemeinsame wissenschaftliche Veranstaltungen gibt, und er fordert die Institute auf: "Bitte benennen Sie die abteilungsübergreifenden Forschungsprojekte und geben Sie an, welche Abteilungen oder Organisationseinheiten daran beteiligt sind."¹⁷

Offensichtlich verleitet diese Formulierung das eine oder andere Institut dazu, abteilungsübergreifende Kooperationen für einen Selbstzweck und viele abteilungsübergreifende Projekte für einen positiven Leistungsnachweis zu halten. Ähnlich verhält es sich mit der abteilungsübergreifenden Forschungsplanung - in einem Fall hat der Wissenschaftsrat angemahnt, daß mehrjährige Forschungsplanungen Beschreibungen zukünftiger Arbeiten sein sollten, aber keine Festlegungen, die eine innere Dynamik nicht zulassen.

Diese Erfahrungen zeigen, wie wichtig es ist, den Instituten den Interpretationsrahmen zu vermitteln, vor dem ihre Antworten erhoben werden. In den Evaluationen sollte es nicht darum gehen, zentrale Strukturmerkmale nur abzufragen, sondern der institutsinterne Diskussions- und Gestaltungsprozeß sollte zur Kenntnis genommen, kritisch bewertet und gegebenenfalls beratend unterstützt werden. Die Frage nach abteilungsübergreifenden Kooperationen sollte mit diesem Ziel um die Aufforderung ergänzt werden, die Prinzipien der Abteilungsgliederung, ihre Vor- und Nachteile im Hinblick auf die Bearbeitung der Forschungsaufgaben sowie die beabsichtigten Strukturentwicklungen darzustellen.

Es ist eine wichtige Aufgabe von Evaluationen, zu überprüfen, ob die Organisationsstruktur noch den aktuellen Erfordernissen angemessen ist oder ob sie verändert werden sollte, und dementsprechende Vorschläge zu unterbreiten. Bisher ist es sehr schwierig, einmal bestehende Abteilungen zu schließen oder mit anderen zusammenzulegen, da davon immer auch persönliche und fachliche Interessen berührt werden. Um eventuellen Konflikten aus dem Wege zu gehen, werden Neuorganisationen offensichtlich gerne dann vorgenommen, wenn Abteilungsleiter ausscheiden. Im Rahmen von Evaluationen dagegen können Erörterungen über den Abteilungszuschnitt zu einem routineförmig angelegten Prozeß werden und diese Debatten versachlichen.

¹⁷ Teil B (Spezifische Fragen), Abs. II. Zur Organisation, Planung und Bewertung der Tätigkeiten, Frage 12 des "Fragebogen für die Bewertung der Forschungseinrichtungen und Museen der Blauen Liste".

2. *Die Organisation der "Dienstleistungen"*. Die Beratungs- und Dienstleistungen werden unterschiedlich organisiert: In einigen Instituten sind einzelne Personen oder kleine Abteilungen damit beauftragt, in anderen Einrichtungen werden die Beratungs-, Informations- und Weiterbildungsaktivitäten von MitarbeiterInnen aller Abteilungen wahrgenommen. Die Belastung mit diesen Aufgaben ist von Institut zu Institut sehr unterschiedlich, und auch innerhalb der Institute gibt es große Unterschiede zwischen den einzelnen MitarbeiterInnen. Dabei kann man nach Status und Aufgaben differenzieren: Leitende WissenschaftlerInnen scheinen häufiger die offizielle Außendarstellung der Einrichtungen zu übernehmen, während sich die Angehörigen des "Mittelbaus" als wissenschaftliche Spezialisten um Einzelfragen kümmern. Für einige MitarbeiterInnen, deren Beratungs- oder Weiterbildungsaktivitäten auf ihnen zurechenbaren Forschungsergebnissen basieren, besteht dabei durchaus die Möglichkeit des Reputationsgewinns, für andere dagegen, denen beispielsweise die Öffentlichkeitsarbeit eines Institutes übertragen wurde, bedeutet sie eher eine zusätzliche Belastung.

Auffällig ist, daß der Stellenwert von Beratungs- und Dienstleistungen in allen Instituten hoch ist, jedoch handelt es sich gleichzeitig aus der Sicht vieler MitarbeiterInnen dabei um zeitaufwendige Aktivitäten, die von der "eentlichen" Forschungsarbeit abhalten und nicht in ähnlichem Maße anerkannt werden. Evaluationen können daher in zweifacher Hinsicht wichtig sein: Wenn im Rahmen von Evaluationen auch die Beratungen und Dienstleistungen der Institute als eigener Leistungsbereich begutachtet, unterstützt und gefördert werden, wird damit zugleich eine wichtige Voraussetzung geschaffen, diese Aufgaben stärker zu professionalisieren, als es derzeit der Fall ist.

3. *Die internen Steuerungsgremien*. Die Evaluation durch den Wissenschaftsrat hat in zweifacher Hinsicht die institutsinternen Diskussionen über angemessene Steuerungsgremien und -verfahren beeinflußt: Zum einen hat die Vorbereitung auf die Evaluation manchen Instituten deutlich gemacht, daß die internen Kommunikations- und Informationsprozesse verbessert werden könnten, zum anderen hat der Wissenschaftsrat selber einige (wenige) Merkmale kritisiert. Die Gutachtergruppen haben den Aufgaben, der Zusammensetzung und dem Zusammenwirken der Gremien unterschiedliche Aufmerksamkeit gewidmet. Die Stellungnahmen sind überwiegend durch Zurückhaltung gekennzeichnet, nur in einem Fall werden nachdrücklich Änderungen - insbesondere eine stärkere Position des Präsidenten - verlangt.

Dagegen haben sich unsere GesprächspartnerInnen eher (selbst)kritisch geäußert. Sie sind insbesondere auf folgende Aspekte eingegangen:

- Ein häufiger Kritikpunkt ist die Zusammensetzung der Gremien. Während die leitenden WissenschaftlerInnen - beispielsweise die AbteilungsleiterInnen - in der Regel qua Funktion Mitglieder in den "Abteilungsleiterkonferenzen" oder im "Wissenschaftlichen Rat" sind, ist die Zahl der Wissenschaftlichen MitarbeiterInnen sowie die Teilnahme und der Status des Betriebs- bzw. Personalrates und der Gleichstellungsbeauftragten häufig strittig.
- In manchen Instituten komme es zu einer Überfrachtung der Gremien mit administrativen und/oder Routineangelegenheiten, die, so einige Interviewpartner, schon aus Zeitgründen eine intensive Erörterung forschungsstrategischer oder forschungspolitischer Probleme kaum möglich mache. Zum Teil hängt die Überlastung mit der Größe der Einrichtung und den entsprechend anfallenden Koordinierungs- und Steuerungsaufgaben zusammen, zum Teil ist eine verstärkte Unterstützung dann gefordert, wenn ein Institutsdirektor auch für den administrativen Geschäftsbereich verantwortlich ist.
- Neben den satzungsmäßig verankerten Beratungsgremien wie dem "Wissenschaftlichen Rat" oder der "Abteilungsleiterkonferenz" gibt es in vielen Einrichtungen weitere, zumeist informelle Gremien. Sie dienen beispielsweise der Steuerung abteilungsübergreifender Kooperationen, der Kontrolle und Qualitätssicherung wissenschaftlicher Publikationen oder der Forschungsplanung und -koordination. Typische Probleme scheinen die unregelmäßigen Zugangsmöglichkeiten zu diesen Gremien, die unklaren Entscheidungs- und Durchsetzungsmöglichkeiten sowie der hohe Abstimmungs- und Informationsaufwand mit anderen Personen oder Gremien (deren Kompetenzen zudem teilweise rechtlich legitimiert sind) zu sein.

Auffallend ist, daß sich die verschiedenen Gremien der Institute im Einzelnen erheblich unterscheiden, was nicht zuletzt mit dem Alter der Einrichtungen und den Folgen unterschiedlicher Gestaltungsversuche zusammenhängt. Dennoch mißt ein großer Teil der von uns Befragten der zukünftigen Aufgabe einer Reorganisation dieser Gremien hohe Bedeutung bei.

4. *Nachwuchsförderung*. Der Wissenschaftsrat hat schon 1993 in seinen Empfehlungen zur Neuordnung der Blauen Liste dargelegt, daß von den Blauen Liste-Instituten eine Beteiligung an der Qualifizierung des wissenschaftlichen Nachwuchses erwartet wird. Er wies darauf hin, daß von den Nachwuchswissenschaftlern wichtige Impulse für die Forschung ausgehen und die Doktoranden zur Flexibilität der Institute beitragen (vgl. Wissenschaftsrat 1993, S. 23 ff.). Auch in der derzeit laufenden Evaluation der Blauen Liste spielt die Förderung des Nachwuchses eine wichtige Rolle. Wir haben daher unsere GesprächspartnerInnen - im Rahmen des großen Komplexes "Personalpolitik", auf den hier nicht ausführlich eingegangen werden kann - zur Promotionsförderung befragt.

Unabhängig davon, daß die interviewten DoktorandInnen mit ihrer Betreuung überwiegend zufrieden sind, benennen die wissenschaftlichen MitarbeiterInnen und die leitenden WissenschaftlerInnen mehrere, aus ihrer Sicht substantielle Probleme der Nachwuchsförderung. Eine Auswahl soll hier kurz skizziert werden:

- Es gebe erhebliche Probleme bei der Rekrutierung wissenschaftlichen Nachwuchses, da man - so die Argumente eines Institutes - den erfolgreichen PromovendInnen angesichts der wenigen frei werdenden Stellen im Hause kaum Hoffnungen auf einen Anschlußvertrag machen könne. Da die Einrichtung auf dem Forschungsfeld singulär tätig sei, falle um so schwerer ins Gewicht, daß die entsprechenden Lehrstühle an den Hochschulen nicht sehr zahlreich seien und zudem noch abgebaut würden.
- Ein weiteres Rekrutierungsproblem besteht für einige Institute der neuen Bundesländer in der Nähe von Berlin, da die Berliner Forschungseinrichtungen und Hochschulen im Unterschied zu den Brandenburger Instituten nach West-Tarif bezahlen können. Außerdem sei die Zahl der Studierenden in naturwissenschaftlich-technischen Fächern erheblich gesunken, was sich bereits in der kleiner werdenden Zahl von qualifizierten BewerberInnen um Promotionsstellen zeige.
- Während einige Institute offenkundig in Antizipation der Erwartungen des Wissenschaftsrates die Anzahl der Promovierenden in ihrer Einrichtung erhöht und damit auch positive Erfahrungen gemacht haben, wurde von einem Institut die Nachwuchsförderung als institutionelle Aufgabe erheblich in Zweifel gezogen. Als anwendungsorientiertes Institut müsse man vor allem erfahrene WissenschaftlerInnen

rekrutieren, die Einsatz- und Betreuungsmöglichkeiten für den wissenschaftlichen Nachwuchs seien eher gering.

Die Konzentration unser GesprächspartnerInnen auf das Rekrutierungsproblem ist auffallend und zeigt, wie unterschiedlich die Voraussetzungen der Institute zur Nachwuchsförderung sind. Die Attraktivität der Institute für DoktorandInnen hängt unter anderem mit den erwarteten Chancen auf dem akademischen Arbeitsmarkt und außerwissenschaftlichen Alternativen, mit disziplinär unterschiedlichen Karrieremustern, regionaler Konkurrenz zwischen den Instituten und nicht zuletzt damit zusammen, ob die Betriebsförmigkeit der Forschung (kurze Projektlaufzeiten, feste Termine etc.) eine Integration von Promovierenden erlaubt. Diese Gesichtspunkte müssen bei Evaluationen berücksichtigt werden, und wenn es zwischen den Instituten erhebliche Unterschiede in der Zahl der DoktorandInnen gibt, werden die Gutachter die schwierige Aufgabe der Beurteilung haben, ob die Unterschiede auf institutsspezifische Besonderheiten oder eine eher zögernde Herangehensweise an die Promotionsförderung zurückzuführen sind. Im übrigen sagt die Zahl der an einem Institut tätigen DoktorandInnen selbstredend wenig über Qualität der Nachwuchsförderung aus - dazu müssen überzeugende Kriterien erst noch entwickelt werden.

Abschließend sei festgehalten, daß die Nachwuchsförderung nur einen kleinen Ausschnitt der personalpolitischen Aufgaben darstellt, und neben der Personalpolitik sind die Institute mit einer ganzen Reihe organisatorischer und struktureller Probleme konfrontiert, die zum Teil bereits genannt wurden: Dazu gehören Fragen der Aufbauorganisation und der interdisziplinären Kooperation, der Ausbau von internen Kommunikations- und Informationsstrukturen und die Gestaltung der Umweltbeziehungen. Die Bewältigung bisher nicht erwähnter Aufgaben wie die Suche nach erfolgreichen Drittmittelstrategien, die Frauenförderung oder die Forschungsk Kooperation mit externen Partnern gehört ebenfalls zu den Anforderungen an die Einrichtungen. Viele unserer GesprächspartnerInnen waren an Informationen und Anregungen interessiert, wie die Gestaltung dieser Aufgaben bewerkstelligt werden kann, und äußerten im Einzelfall Unterstützungs- und Beratungsbedarf.

Sicherlich wird der Zusammenschluß der Blaue Liste-Einrichtungen in der WGL einen Erfahrungsaustausch - etwa durch die Kooperation fachlich verwandter Institute in den Sektionen - befördern. Darüber hinaus sehen wir große Chancen, den erwähnten Bedarf

in den Evaluationen nutzbringend aufzugreifen. Insbesondere im Kontext der Begutachtung durch die Beiräte, die auch mit den fachlichen Besonderheiten der Institute vertraut sind, besteht die Möglichkeit einer intensiven Erörterung organisatorischer und struktureller Probleme. Das würde allerdings darauf hinauslaufen, nicht nur die Forschungsleistungen, sondern auch die sogenannten "Rahmenbedingungen" der Forschung zum Gegenstand von Evaluationen zu machen, und es würde bedeuten, den Aspekt der "Beratung" neben der "Bewertung" zu betonen - darauf wird später zurückzukommen sein.

3. Die Suche nach angemessenen Indikatoren

Für die Evaluation der Blaue Liste-Institute sind mehrere Konzepte von Bedeutung: Bund und Länder haben sich in der "Rahmenvereinbarung Forschungsförderung" darauf verständigt, regelmäßig die "überregionale Bedeutung" und das "gesamtsstaatliche wissenschaftspolitische Interesse" zu überprüfen. Als weitere wichtige Voraussetzung für die Finanzierung der Blaue Liste-Einrichtungen durch Bund und Länder überprüft der Wissenschaftsrat auch das Leistungsmerkmal "Qualität der Forschung".

Angesichts der Vielfältigkeit und Unterschiedlichkeit der Aufgaben von Blaue Liste-Instituten fragt sich, was unter "Qualität der Forschung" zu verstehen ist, wie Qualität gemessen oder bewertet werden kann und ob dieses Konzept in der Lage ist, die gesamte Bandbreite der Leistungen, die von diesen Einrichtungen erbracht werden, zu erfassen. Welche Merkmale machen die "Qualität" von Forschung aus, und warum werden nicht auch die "Nützlichkeit" oder die "Relevanz" von Forschung evaluiert? Welche Funktion sollten Wissenschaftsindikatoren haben? Mit welchen Instrumenten können auch Informations- und Beratungsaufgaben berücksichtigt werden?

"Qualität" ist ein schwieriger, weil zwar viel verwendeter, aber nebulös bleibender Begriff. Weil Qualität von Forschung schwer zu definieren und festzustellen ist, "... bleibt nur die verbale Umschreibung oder das Qualitätsurteil, vom Fachmann (oder Komitees) abgegeben und vom kundigen Leser ausgewertet." (Alewell 1993, S. 161 f.). Beispielsweise werden Projektanträge durch Fachkollegen begutachtet beziehungsweise in der Form von Peer Review-Verfahren bewertet, und "Peers" geben Gutachten in

Berufungsverfahren oder bei der Veröffentlichung von Manuskripten in Fachzeitschriften ab.

Peer Review ist allerdings keineswegs unumstritten: "Peer Review ist nach wie vor das am häufigsten verwandte Verfahren zur Bewertung im Wissenschaftsbereich, und es ist zugleich wohl auch das am heftigsten und häufigsten attackierte Verfahren" (Hornbostel 1997, S. 196). Zu den häufigsten Beanstandungen gehört, daß die Gutachter nicht unabhängig und ihre Urteile subjektiv geprägt seien. Mit der Expansion des Wissenschaftssystems und der Flut von neuen Publikationen, Zeitschriften und Forschungsanträgen scheint darüber hinaus die Belastbarkeit des Gutachtersystems an Grenzen zu stoßen.

Unter anderem aufgrund der Schwächen von Peer Review-Verfahren wird seit mehreren Jahren über *Wissenschaftsindikatoren* diskutiert, mit deren Hilfe die Qualität von individuellen oder institutionellen Leistungen bewertet werden soll. Als Indikator für Forschungsleistungen sind insbesondere Publikationen, Angaben über die Zahl der zitierten Veröffentlichungen und Drittmiteleinwerbungen von Bedeutung. Zwar werden die Wissenschaftsindikatoren häufig für "objektiver" gehalten als Begutachtungen von Fachkollegen, jedoch kommen die Indikatoren keineswegs ohne Elemente von Peer Review aus, wie ein kurzer Blick auf zwei Indikatoren zeigt.

Ein wichtiger Indikator, der sowohl für die Begutachtung von Personen wie von Institutionen eingesetzt wird, ist die Zahl der *Publikationen*. Um dem Einwand zu begegnen, daß ein solcher Indikator die Produktivität von WissenschaftlerInnen mit der Qualität der Forschungsergebnisse verwechselt, werden nicht alle Publikationstypen bewertet. Selbstredend scheidet die "graue" Literatur aus, und Publikationen in referierten Zeitschriften haben ein größeres Gewicht als Veröffentlichungen in anderen Medien. Zwar ist das Argument einleuchtend, daß für eine Veröffentlichung in einer international erscheinenden und referierten Zeitschrift eine weit höhere Hürde genommen werden muß als für eine Veröffentlichung in einer Hauszeitschrift, dennoch haben sich durch die Urteile der "Referees" indirekt Peer Review-Verfahren Geltung verschafft.

Auch die Höhe der eingeworbenen *Drittmittel* ist ein wichtiger Indikator. Es wird argumentiert, daß die Bewilligung von Drittmitteln auf der Grundlage einer "Qualitätskontrolle" von Projektanträgen basiert, die sich zudem noch in Konkurrenz zu anderen

Anträgen durchsetzen müssen. Andererseits kann die Höhe der Drittmittel jedoch auch ein Hinweis darauf sein, daß sich externe Interessen geltend machen oder ein Institut in eine enge Auftraggeber-Auftragnehmer-Beziehung geraten ist. Um nur solche Drittmittel zu bewerten, die tatsächlich das Ergebnis einer Selektion unter wissenschaftlichen Kriterien sind, wird daher in manchen Verfahren nicht nur die *Höhe* der Drittmittel berücksichtigt, sondern auch ermittelt, welche *Herkunft* die Drittmittel haben. Beispielsweise legt der Wissenschaftsrat bei seiner Evaluation der Blaue Liste-Institute besonderen Wert auf Drittmittel, die die Institute von der Deutschen Forschungsgemeinschaft (DFG) eingeworben haben. Hintergrund ist vermutlich, daß für DFG-Anträge erhebliche Vorarbeiten geleistet werden müssen, die DFG für ihre ausgesprochen strenge Begutachtung bekannt ist und insofern die Zuwendung von DFG-Mitteln als Ausweis für herausragende Leistungen gilt. Die große Bedeutung von Peer Review ist also auch bei diesem Indikator offenkundig.

Mittlerweile hat sich die Erkenntnis durchgesetzt, daß die Aussagekraft der Wissenschaftsindikatoren begrenzt ist: Sie können im Rahmen von Peer Review-Verfahren ergänzende quantitative Informationen zur Verfügung stellen und die Gutachter orientieren, sie müssen allerdings ihrerseits interpretiert werden. Es gibt eine breite Diskussion über die Verlässlichkeit und Gültigkeit von Wissenschaftsindikatoren, die hier nicht ausführlich dargestellt werden soll. Wir wollen in unserer Studie insbesondere auf die folgenden drei Aspekte eingehen:

Erstens sollen die erwähnten Wissenschaftsindikatoren offenkundig besonders herausragende "Spitzenleistungen" identifizieren. Die Voraussetzungen dazu sind jedoch in den Blaue Liste-Instituten sehr ungleich verteilt. Beispielsweise kann man von einem grundlagenorientierten Institut erwarten, daß es eine gewisse Zahl von Publikationen in englischsprachigen, referierten Zeitschriften nachweist, aber nicht unbedingt von einem stark anwendungsorientierten Institut, das einen eingeschränkten nationalen Markt "bedient". Andererseits wird man von diesem Institut aber erwarten können, daß es seine Forschungsergebnisse beispielsweise gezielt in solchen Medien veröffentlicht, die bestimmte Adressaten - wie zum Beispiel die mittelständische Industrie - erreicht. Daher müssen die gültigen Indikatoren unter Berücksichtigung des jeweiligen Aufgabenprofils eines Institutes definiert und differenziert werden. Zu klären ist, welche Besonderheiten dabei von Bedeutung sind.

Zweitens können durch die Wissenschaftsindikatoren diejenigen Leistungen und Angebote, welche von den WGL-Instituten neben den Forschungsergebnissen zur Verfügung gestellt werden, nicht erfaßt werden. In unserer Studie soll untersucht werden, wie diese Leistungen - zum Beispiel Beratungen oder von den Einrichtungen angebotene Schulungen und Weiterbildungsmaßnahmen - charakterisiert werden können und ob eine Bewertung durch quantitative oder qualitative Verfahren möglich ist.

Drittens schließlich gehen wir davon aus, daß die eingesetzten Indikatoren erhebliche Rückwirkungen auf die Forschungspraxis in den Instituten haben. Beispielsweise ist bekannt, daß die Verwendung von Indikatoren, die auf der Zahl der Publikationen basieren, dazu führen kann, daß Forschungsergebnisse "gestückelt" publiziert werden oder daß WissenschaftlerInnen versuchen, ihre Ergebnisse unter wechselnden Überschriften und mit leichten Varianten in mehreren Zeitschriften unterzubringen. Indikatoren wirken also als ein Steuerungs- und Anreizinstrument - diese Funktion ist erwünscht, sie kann aber auch eher negative Folgen haben. Im Rahmen unserer Studie haben wir versucht, auch solche nicht intendierten Folgen zu ermitteln.

3.1 Die Dominanz "akademischer" Indikatoren

In vielen Forschungsevaluationen spielen Publikationsanalysen, Zitationsanalysen und Untersuchungen des Drittmittelaufkommens eine herausragende Rolle. Auch im Rahmen der derzeitigen Evaluation durch den Wissenschaftsrat haben diese Wissenschaftsindikatoren eine große Bedeutung. Zwar wird von Zitationsanalysen abgesehen, aber wichtige Indikatoren sind die Höhe der Publikationen in referierten Zeitschriften und die Höhe der von der DFG eingeworbenen Drittmittel.

Wir haben daher im Rahmen dieser Studie die Chance gesehen, aufgrund von Befragungen und Beobachtungen in mehreren Instituten Hinweise auf die Validität der Indikatoren "DFG-Mittel" und "Publikationen in referierten Journals" zu gewinnen. Einige wichtige Befunde dazu lassen sich - nicht abschließend - wie folgt zusammenfassen:

- Wenn eine Forschungseinrichtung nicht nur wissenschaftsinterne Adressaten hat, sondern die Generierung von Forschungsfragen auch unter wissenschaftsexternen Aspekten vorgenommen wird, sind Publikationen in referierten Journals nur ein

unvollständiger Leistungsindikator. Man könnte diesen Sachverhalt am Beispiel des *Institut B* verdeutlichen: Dieses Institut erforscht das Lehr-Lern-Verhalten im mathematisch-naturwissenschaftlichen Unterricht. Dazu werden international vergleichende und grundlagenorientierte Studien durchgeführt, die gute Chancen haben, in referierten Journals publiziert zu werden; gleichzeitig entwickelt das *Institut B* didaktische Konzepte für den naturwissenschaftlichen Unterricht und erarbeitet Materialien, die in Weiterbildungsveranstaltungen für LehrerInnen eingesetzt werden. Für beide Aufgaben müssen unterschiedliche Publikationstypen erstellt werden - wenn eine Publikationsform vernachlässigt wird, könnte das zugleich die Vernachlässigung eines wichtigen Institutsziels bedeuten.

- Für den Publikationserfolg in referierten Journals ist aber nicht nur eine Orientierung an wissenschaftsintern generierten Fragestellungen wichtig, sondern auch die Struktur des Forschungsfeldes. Je nach der Struktur dieses Feldes kann der Zugang zu referierten Zeitschriften schwierig sein. Auch diese Aussage kann am Beispiel des *Institut B* verdeutlicht werden: Da es keinen gemeinsamen Publikationsmarkt für die Pädagogik und Psychologie einerseits und die Fachdidaktiken andererseits gibt, ist der Zugang zu referierten Zeitschriften zwar möglich, aber aufwendig. Im Gegenzug gibt das *Institut B* Monographien und Sammelbände in einer hauseigenen Reihe heraus, die forschungsstrategisch als Bündelung von interdisziplinären Forschungsfragen eine wichtige Rolle spielt.
- Viele referierte Zeitschriften sind englischsprachig und international verbreitet - darin werden nur solche Forschungsergebnisse gewisse Publikationschancen haben, die sich an der sogenannten "Forschungsfront" bewegen. Spezifisch nationale oder gar regionale Fragestellungen werden aus Sicht der Herausgeber kaum von Interesse sein. Gerade darauf beziehen sich aber wichtige Forschungsfragen der von uns untersuchten WGL-Institute - sie beziehen aus ihrer Konzentration auf nationale und/oder regionale Probleme ihre besondere Legitimation. Auf die besondere Situation der Hochschuldidaktik in den Naturwissenschaften, mit der sich das *Institut B* auseinandersetzen muß, wurde bereits hingewiesen; das *Institut C* hat - wie auch der Wissenschaftsrat betonte - sowohl hinsichtlich der Forschungs- wie bei den Beratungsaufgaben - national einzigartige Aufgaben wahrzunehmen; die Forschungsfragen des *Institut A* sind in großen Teilen national- oder sogar regionalspezifisch, und auch das *Institut D* bearbeitet Probleme, die letztlich spezifische, geologisch einzig-

artige Hintergründe haben. Auch das *Institut E* verdankt seine Existenz nationalen Besonderheiten: Aufgrund der großen technologischen Bedeutung für verschiedene innovative Forschungsgebiete beteiligt sich das Institut - mit international beachteten Ergebnissen - an der Grundlagenforschung. Die dabei gewonnenen Erkenntnisse dienen - und darin unterscheidet es sich etwa von Max-Planck-Instituten - der zentralen Aufgabe, für andere Forschungsinstitute, für universitäre Forschungseinrichtungen und vor allem für die Industrie Kristalle herzustellen, die jeweils singuläre Anforderungen erfüllen müssen. Ein Übermaß an Publikationen in referierten Journals kann daher keineswegs immer ein "Gütesiegel" sein - im Gegenteil sehen sich die Institute mit einer aufwendigen Erhöhung dieser für die scientific community interessanten Publikationen der Gefahr ausgesetzt, andere wichtige Aufgaben zu vernachlässigen.

- Mit der Veröffentlichung in referierten Zeitschriften ist immer ein gewisser Zeitverlust verbunden, da die jeweilige Zeitschrift zunächst Gutachter für die eingereichten Arbeiten suchen wird und die Gutachter eine positive Empfehlung häufig mit einzelnen Überarbeitungsaufträgen verbinden. Auf manchen Forschungsfeldern ist jedoch eine rasche Publikation oder eine hohe Aktualität der Forschungsergebnisse wichtiger als das Renommee einer Zeitschrift.
- Hinsichtlich der DFG-Mittel gelten schon bekannte Argumente: Die Akquisition von DFG-Mitteln darf kein Selbstzweck werden, sondern es muß eine Relation zur Relevanz der Grundlagenforschung in der jeweiligen Forschungseinrichtung erkennbar sein. Eine besondere Berücksichtigung von DFG-Mitteln wäre manchen Einrichtungen und ihren jeweiligen Zielen nur wenig angemessen - beispielsweise wäre dem *Institut D*, einer auch umsetzungsorientierten Einrichtung, damit wenig gedient. Dagegen kann in diesem Fall die Vielfalt der Drittmittelgeber - von Bundes- und Länderministerien bis zu EU-Mitteln - ein Hinweis auf die gute Kooperation mit Akteuren sein, deren Interessen durchaus unterschiedlich sind. Insofern ist diese "Durchmischung" ein Indikator dafür, daß das *Institut D* einem Teilziel innerhalb seines Leistungsspektrums - die ökologische Umgestaltung von Landschaften auch als einen Moderationsprozeß zwischen unterschiedlichen Akteuren zu realisieren - näher gekommen ist.

- Viele InterviewpartnerInnen sind außerdem an der Einwerbung von DFG-Mitteln nur wenig interessiert. Es ist sehr zeit- und arbeitsaufwendig, DFG-Mittel einzuwerben, und sie werden für problematisch gehalten, da DFG-Projekte meistens im Sach- und Reisemittelbereich kofinanziert werden müssen. Wichtiger noch ist, daß DFG-Mittel nach disziplinären Gesichtspunkten vergeben werden. Nun soll jedoch eine wesentliche Aufgabe von WGL-Instituten gerade in der interdisziplinären Kooperation bestehen - diese wird durch DFG-Mittel nicht nur nicht "belohnt", sondern im Gegenteil werden sogar disziplinäre Orientierungen verstärkt.

Aus dieser Argumentation ist keineswegs zu schlußfolgern, daß die Einwerbung von DFG-Mitteln oder Publikationen in referierten Zeitschriften nicht als Bewertungsmaßstab gelten sollten - nur sind sie im Kontext des gesamten Aufgaben- und Leistungsspektrums der Institute zu interpretieren und zu gewichten. Um jene Leistungen zu bewerten, die stärker in den anwendungsorientierten Bereich fallen oder zu den außerwissenschaftlichen Aufgaben gehören, sollten darüber hinaus auch für diese Tätigkeitsgebiete Leistungskriterien entwickelt werden. So ließe sich ein wichtiger Aufgabenbereich adäquat würdigen, und man könnte die konventionellen, akademisch-grundlagenorientierten Indikatoren um andere Maßstäbe ergänzen. Eine Quantifizierung dieser Tätigkeiten scheint jedoch nach den Aussagen unserer InterviewpartnerInnen nur sehr begrenzt möglich zu sein. Auf die Art der verschiedenen Leistungen wurde bereits eingegangen: Es handelt sich beispielsweise um die Durchführung von Weiterbildungsmaßnahmen, um die Beantwortung von schriftlichen und telefonischen Anfragen, um Politikberatung, Öffentlichkeitsarbeit und um Kontakte zu den Medien. Ungewöhnlich sind die erwähnten Leistungen des *Institut E*, das Kristalle züchtet und diese anderen Forschungseinrichtungen und der Industrie zur Verfügung stellt. Mehrere MitarbeiterInnen meinen daher, daß die zu bewertenden "Produkte" des Hauses nicht die Publikationen, sondern die Qualität der Kristalle sein müsse. Das *Institut E* nimmt aber wohl einen Sonderstatus ein, von den anderen Instituten wurden vor allem hinsichtlich der Öffentlichkeitsarbeit Ergänzungen der Indikatoren vorgeschlagen. So sollten auch Veröffentlichungen in populären Zeitungen und Zeitschriften sowie Fernseh- und Radiobeiträge erfaßt werden. Dagegen würde z.B. der Versuch, die telefonischen Beratungen zu quantifizieren, indem die Anfragen gezählt werden, nur zu einer abstrusen Bürokratisierung des Alltags führen. Außerdem würde, und damit sind wir wieder bei einem zentralen Gegenargument, die Quantifizierung dieser Leistungen mit Aussagen über ihre Qualität verwechselt.

Zuletzt soll auf die erwähnten "steuernden" Effekte der Evaluation eingegangen werden. Im Rahmen der Evaluation durch die Bewertungsgruppen des Wissenschaftsrates wurden und werden neben den Wissenschaftsindikatoren auch Kriterien verwendet, mit denen die strukturell-organisatorischen Voraussetzungen für "gute Forschung" bewertet werden sollen. Beispielsweise wird eine Befristungsquote von 30% bis 50% erwartet und besonderer Wert auf die Förderung des wissenschaftlichen Nachwuchses gelegt. Selbstredend hatten alle Institute Kenntnis von den eingesetzten Bewertungskriterien; sie haben sich - zum Teil lange - im Vorfeld auf die Evaluation vorbereitet und entsprechende Aktivitäten, beispielsweise zur Beschäftigung von PromovendInnen oder im Rahmen der Publikationspolitik, entfaltet.

Das heißt aber, daß die verwendeten Indikatoren nicht nur "messen" und "bewerten" und ansonsten keinen Einfluß auf die Forschungspraxis haben. Sie besitzen im Gegenteil ein erhebliches Steuerungspotential, das zum Teil bereits erwähnt wurde:

- Die indirekte "Belohnung" von bestimmten Publikationstypen - den Veröffentlichungen in internationalen, referierten Journals - veranlaßt eine stärkere Orientierung an wissenschaftsintern generierten Fragestellungen,
- zu Lasten von nationalen oder regionalen Fragestellungen wird eine Ausrichtung am internationalen "mainstream" vorgenommen, da sie einen höheren Publikationserfolg in renommierten Zeitschriften verspricht,
- die fehlenden Quantifizierungsmöglichkeiten für Beratungs- und Dienstleistungen führen dazu, daß im Gegenzug die viel stärker "sichtbaren" Forschungsleistungen ausgebaut werden,
- die Drittmittelakquise konzentriert sich aus zwei Gründen in erheblichem Maße auf grundlagenorientierte DFG-Mittel: Sie werden erstens vom Wissenschaftsrat besonders hoch bewertet. Zweitens sind die WGL-Institute dazu verpflichtet worden, 2,5% ihrer Grundfinanzierung an die DFG abzuführen - diese Mittel können im Gegenzug über die Einwerbung von Drittmitteln bei der DFG wieder "zurückgeholt" werden.¹⁸

¹⁸ Es handelt sich dabei um eine Initiative des BMBF. Die sogenannte "DFG-Abgabe" ist eine Voraussetzung für die vom BMBF gewünschte wettbewerbsorientierte Vergabe von Projektmitteln der DFG

Möglicherweise sind einzelne Steuerungseffekte der Evaluation intendiert gewesen. In der Tat mag es durchaus sinnvoll sein, nicht nur abzufragen, inwieweit ein Institut in der Lage ist, sich an der internationalen Spitzenforschung zu beteiligen, sondern es damit zugleich zu ermutigen, in der Grundlagenforschung aktiver zu werden. Manche MitarbeiterInnen haben deutlich gemacht, daß sie diese Akzentsetzung des Wissenschaftsrates als anregend empfunden haben, da sie einzelnen Personen oder dem ganzen Haus neue Optionen eröffne.

Andererseits kann gerade durch die Addition mehrerer Indikatoren, die Grundlagenforschung besonders hoch bewerten, also durch einen Bias im Indikatorensatz, eine Gewichtsverlagerung zu Lasten von anwendungsorientierten Forschungen und Dienstleistungen einsetzen. Dadurch würde aber nicht nur langfristig die Verknüpfung unterschiedlicher Forschungstypen gefährdet, die gerade die Besonderheit der Blaue Liste-Institute ausmacht, sondern die Institute würden auch in eine prekäre Nähe zu anderen, ebenfalls grundlagenorientierten Wissenschaftseinrichtungen geraten. Ohne es zu wollen, müßten die Blaue Liste-Institute eine forschungspolitisch gefährliche Konkurrenz auf Feldern aufnehmen, die andere Einrichtungen als latente Bedrohung¹⁹ oder als Einbruch in ihre Domäne begreifen würden.

4. Ziele und Verfahren: Beratung und Bewertung

Die Diskussion über Ziele und Aufgaben, Verfahren, Methoden und Akteure von Evaluationen im Wissenschafts- und Forschungsbereich ist in Deutschland immer noch weitgehend hochschulbezogen. Etabliert sind mittlerweile fast in allen Bundesländern Lehrberichte als ein Bestandteil von qualitätssichernden Elementen. Aber selbst in dieser - im Vergleich zu außeruniversitären Forschungsinstituten -

an die WGL und die Hochschulen. Diese beiden Akteure stehen allerdings dem Vorhaben des BMBF kritisch gegenüber: Während die Hochschulen vermuten, daß der Wettbewerb angesichts höher spezialisierter und besser ausgestatteter Forschungseinrichtungen zu ihren Lasten ausgehen wird, sorgt sich die WGL, daß "... es sich bei der vom BMBF vorgeschlagenen Maßnahme lediglich um eine Umwidmung von Mitteln der WBL an die Hochschulen handelt, denen an anderer Stelle immer wieder Mittel entzogen werden." (Hertel 1998, S. 46)

¹⁹ Die Befürchtungen der Hochschulen, daß ihnen in der WGL ein gewichtiger Konkurrent bei der Vergabe von DFG-Mitteln erwächst, wurden bereits erwähnt. Vermutlich werden die Sorgen nicht geringer, da sich abzeichnet, daß die Drittmittel aus der Wirtschaft zwar insgesamt steigen, jedoch profitieren davon weniger die Hochschulen als die außeruniversitären Forschungseinrichtungen (vgl. Seeling/Sproß 1999).

elaborierten Debatte befindet sich eine systematische (wissenschaftspolitische) Auseinandersetzung über unterschiedliche Zielsetzungen und damit verbundenen Methoden - Transparenz, Rechenschaftslegung, Qualitätsförderung, Leistungsvergleich, leistungsbezogene Mittelvergabe, um nur einige zu nennen - erst am Anfang (vgl. Barz, Carstensen, Reissert 1997, S. 39 f.). Dabei erfordern gerade die inhaltlichen und institutionellen Veränderungen der Wissenschafts- und Forschungslandschaft eine systematische Reflexion und auch Anpassung von Evaluierungsverfahren an neue Anforderungen und Problemlagen.

Daß Qualität und Qualitätsbeurteilung - hier von Forschung - im Zentrum von Evaluationen steht, ist nun keine neue Erkenntnis. Bedeutender sind hier die unterschiedlichen Modelle und Zielvorstellungen von Evaluierungen hinsichtlich qualitätskontrollierender, -sichernder und -fördernder Elemente mit den entsprechenden unterschiedlichen wissenschaftspolitischen Implikationen. Vroeijenstijn hält aus den Erfahrungen der Evaluierungspraxis der Hochschulen in den Niederlanden Qualitätskontrolle für das Wissenschafts- und Forschungssystem für nicht angemessen, jedoch Qualitätssicherung und -förderung (Vroeijenstijn 1995). "Quality control is inherently punitive, imposing sanctions for inadequacy, but at the same time it implies at once a minimum acceptable level has been reached no further effort for improvement is needed." (Lynton 1988; zitiert in Vroeijenstijn 1995, S. 8).

Auch die Deutung und Interpretation des Qualitätsbegriffs ist je nach Akteur und spezifischen Interessenskonstellationen unterschiedlich (und der Bedarf nach systematischer Aufarbeitung offensichtlich). Selbst bei so zentralen Aspekten von Qualitätsgrößen im Hochschulbereich wie "Input" (z.B. student intake, selection, academic staff), "Process" (z.B. aims/goals, educational organisation, counselling) und "Output" (z.B. pass/fail rate, the graduate) bestehen seitens der involvierten Akteure (Studenten, wissenschaftliches Personal, Hochschulverwaltung und Wissenschaftsministerien) unterschiedliche Interessen, Gewichtungen und Orientierungen (ebenda, S. 14).

In bezug auf die Leistungen und "Produkte" von Hochschulen kann nach Harvey und Green zwischen fünf Kategorien unterschieden werden: "Quality as exceptional, quality as perfection or consistency, quality as fitness for purpose, quality as value for money, quality as transformation. This grouping spans the universe from traditional

notions of academic quality as excellence ('exceptional') to the recent insights of 'zero defects' ('perfection'), mission orientation and consumer orientation ('fitness for purpose'), and finally, in the transformational notion of quality, to the question of what higher education is about" (zitiert nach Westerheijden, Brennan, Maassen 1994, S. 16/17).²⁰

Mit Blick auf Qualitätssicherungsprozesse wird zwischen zwei Evaluationstypen differenziert: Der *summativen* und der *formativen*. Diese von Scriven 1967 eingeführte Unterscheidung, die sich nicht nur auf den Wissenschaftsbereich im engeren Sinne, sondern auch beispielsweise auf Programmimplementationen bezieht, fungiert nach wie vor als ein Bezugspunkt, der eine systematische Differenzierung zwischen verschiedenen Evaluationstypen ermöglicht. Während sich der *formative* Ansatz vorrangig auf die Prozessebene (etwa auf laufende Programme) und auf "kurative" Elemente (mit Steuerungs- und Optimierungsfunktion zur Verbesserung von Programmen) konzentriert - hierbei spielen institutsinterne Instrumente der Selbstkontrolle eine besondere Rolle -, steht die Beurteilung und Zertifizierung von Leistungen - also des Outputs - im Mittelpunkt *summativer* Evaluationen (vgl. Grüner 1993). Als Beispiele für den letztgenannten Typ führen Webler u.a. die Hochschulprüfungen des Wissenschaftsrats als Voraussetzung zur Aufnahme dieser Hochschulen in das Hochschulverzeichnis nach dem Hochschulförderungsgesetz oder die Überprüfung von Sonderforschungsbereichen und Forschergruppen der Deutschen Forschungsgemeinschaft auf; für den formativen Typus die Selbstevaluationen der Fachbereiche zum Beispiel zur Erstellung von Lehrberichten oder die Evaluation durch eine Beratungsagentur von Hochschulexperten im Rahmen der Organisationsentwicklung (vgl. Webler 1996, S. 40 ff.).

²⁰ In der - im Vergleich mit der internationalen Entwicklung - ebenfalls verspäteten Debatte in Deutschland über Qualitätsorientierung öffentlicher Dienstleistungen im Zusammenhang mit dem Reformprozeß öffentlicher Verwaltungen findet zur Zeit eine Diskussion über den Qualitätsbegriff statt. Qualität wird nicht mehr vorrangig auf ihre Bedeutung für die Bürger bezogen, sondern gleichzeitig als Parameter für Wettbewerb, für die Steuerung durch Politik und für Verwaltungsmanagement und Personalführung eingesetzt (vgl. Oppen 1995, S. 38/39). Der im Vergleich dazu eher schwerfälliger Diskurs in der Wissenschaft hängt u.a. mit dem "sozialen System Wissenschaft" selber zusammen, da es um Methoden der Einflußnahme auf dieses System geht, die über dessen eigene Mechanismen ("peer review") hinausgehen und nicht mehr von ihr unmittelbar steuerbar sind (vgl. Weingart 1995, S. 2 f.).

Ein anderer Versuch der Kategorisierung ist die Differenzierung zwischen

- (a) *interner unterstützender Beobachtung*, die von der zu bewertenden Einheit nahestehenden Fachleuten durchgeführt wird und die die frühzeitige Identifizierung von Fehlentwicklungen erlaubt - gewährleistet ist ein hoher Grad an Selbststeuerung;
- (b) *interner bewertender Begutachtung*, die zur direkten Entscheidungsfindung oder als Krisenmanagement eingesetzt wird;
- (c) *externer unterstützender Begutachtung*, die relativ selten vorzufinden ist, da externe Evaluationen meistens auf Bewertungen ausgerichtet sind und
- (d) *externer bewertender Begutachtung* etwa durch Ministerien oder andere Zuwendungsgeber, die als Managementinstrumente und vor allem zur Kontrolle dienen (vgl. Felt/Nowotny/Taschwer 1995 S. 239 ff.).

In diesem Zusammenhang wird auf ein Verständnis von Evaluationen als "sozialen Aushandlungsprozessen" und damit zusammenhängend auf die Vielzahl von Methoden, Maßstäben und Kriterien, Perspektiven und Akteursinteressen verwiesen; d.h. eine Favorisierung von offeneren Ansätzen und Konzepten im Vergleich zu standardisierten Verfahren.

Mit Verweis auf die bereits eingeführte Evaluationspraxis an den Hochschulen wäre es sicherlich zu einfach, die Kategorisierung in formative und summative bzw. beratende und bewertende Evaluationsmodelle mit internen und externen Lehrevaluierungen gleichzusetzen; dennoch sind ähnliche strukturierende und zielsetzende Elemente in den zum Teil noch modellartigen Praxen vorzufinden. Webler unterscheidet in diesem Zusammenhang erstens zwischen der Selbstevaluation des Fachbereichs, zweitens der Evaluation des Fachbereichs mit größerem Methodeneinsatz, vergleichend mit anderen Fachbereichen und mit Hilfe von Organisationsberatern aus dem Hochschulbereich, und drittens der Evaluation durch eine Kommission von Fachkollegen (Peer Review) (vgl. Webler 1996, S. 142/143).

Elaborierte Modelle der internen und externen Evaluation findet man vor allem im niederländischen Hochschulsystem, in Deutschland am ehesten noch im Verbund

Norddeutscher Universitäten, die im Rahmen der internen Selbstevaluation Raum für selbstreflexive Kommunikationsprozesse lassen (vgl. Lüthje 1995, S. 148 ff.).²¹

Hinsichtlich außeruniversitärer Forschungseinrichtungen, ihrer spezifischen Aufgaben, Ziele, Organisationstypen, Wissensstrukturen, Forschungstechniken und Bezugsgruppen im Vergleich zu den Hochschulen ist die Diskussion über unterschiedliche Evaluationsmodelle und -verfahren, verbunden mit unterschiedlichen Zielvorstellungen und Zweckbestimmungen, noch nicht so weit entwickelt. Impulse erhielt sie durch die Aktivitäten des Wissenschaftsrats im außeruniversitären Sektor.

1990/91 nahm der Wissenschaftsrat erstmals eine systematische, flächendeckende Evaluation mehrerer Forschungsorganisationen und ihrer Institute vor: Er bewertete - unter anderem - die Institute der Akademie der Wissenschaften der DDR. Innerhalb kurzer Zeit wurde auf der Grundlage dieser Evaluationen entschieden, welche Forschungseinrichtungen aufgelöst und welche Forschergruppen in das bundesrepublikanische Forschungssystem integriert werden sollten. Viele Arbeitsgruppen erhielten eine Überlebenschance im Rahmen der heutigen WGL.

Diese systematische Evaluation durch den Wissenschaftsrat wiederholt sich - auf der Grundlage eines leicht modifizierten Verfahrens - seit 1995. Es ist erklärtes Ziel dieser Bewertungen, die Qualität aller Institute zu überprüfen und durch das Ausscheiden von Einrichtungen die finanziellen Voraussetzungen zu schaffen, andere, leistungsfähigere Institute in die gemeinsame Förderung von Bund und Ländern aufzunehmen. Die Qualitätsüberprüfung soll eine Entscheidung erlauben, ob die Weiterförderung des jeweiligen Instituts im gesamtstaatlichen überregionalen forschungspolitischen Interesse liegt und somit eine Existenz als gemeinsam von Bund und Land finanzierte außeruniversitäre Einrichtung zu rechtfertigen ist.

Im Vordergrund der Wissenschaftsratsevaluationen steht die *Bewertung* von Forschungs- und Dienstleistungen. Darauf basierend werden Empfehlungen zu Fragen der Forschungsleistungen, der Forschungsstruktur und -organisation, zu

²¹ Ähnliche Ziele wie diese Evaluationsmodelle, nämlich als Zweckbestimmung die Selbstvergewisserung des Fachbereichs, Erzeugung von Kommunikations- und Einigungsprozessen als Teil einer Organisationsentwicklung und keine Orientierung auf die Abfassung von Qualitätsurteilen, sondern das Aufzeigen einer Entwicklungsperspektive, verfolgt das Bielefelder Modell zur Evaluation der Lehre als Organisationsberatung durch Hochschullehrer (vgl. Webler 1998, S. 189 ff.).

Kooperationsformen etc. abgegeben. Sie haben insofern eine hohe Relevanz in den Evaluierungsverfahren, da vom Wissenschaftsrat auf die Umsetzung früherer Empfehlungen in den aktuellen Evaluationen direkt Bezug genommen wird. Allgemein den Blauen Liste-Instituten empfohlene qualitätssichernde Maßnahmen wie befristete Besetzungen eines erheblichen Anteils der Stellen des wissenschaftlichen Personals, die Kooperation mit den Hochschulen sowie die internationale Zusammenarbeit gerieten zu impliziten Bewertungskriterien (vgl. Wissenschaftsrat 1993). Insbesondere der Zusammenarbeit mit den Hochschulen wurde entscheidende Bedeutung für die Erhaltung und Verbesserung der Qualität der Forschung außeruniversitärer Forschungseinrichtungen beigemessen (ebenda) und in den Befunden der Evaluationen die Forschungsk Kooperationen, gemeinsame Berufungen, Lehre, wissenschaftlicher Nachwuchs und Wissenschaftlertausch bewertet.

Mit unserer Untersuchung ist zweierlei intendiert:

Erstens interessieren die Zielsetzung und Verfahrenselemente der Wissenschaftsrats-evaluation hinsichtlich der Frage, was dieses Verfahren leistet bzw. leisten kann. Dabei wird der Sichtweise, spezifischen Ausgangssituation und Problemlage von Blauen Liste-Instituten besondere Aufmerksamkeit geschenkt - insbesondere mit Blick auf den Umfang der erhobenen Daten und Informationen, die Vorbereitung, Durchführung und Auswirkungen ("Lerneffekte") auf die Institution, die beteiligten Akteure, die Rezeption und die Umsetzung der Empfehlungen des Wissenschaftsrats.

Zweitens beabsichtigen wir, die im WZB gewonnenen Erfahrungen für die Entwicklung und Erprobung von Beobachtungs- und Evaluationsverfahren zu nutzen und zu erkunden, inwieweit sie sich auch auf Institute mit divergierenden Forschungsprofilen und disziplinären Schwerpunkten anwenden lassen. Dabei kommen unterschiedliche Organisationstypen, Funktionen, Wissensstrukturen, Forschungstechniken und Bezugsgruppen ins Spiel. In diesem Zusammenhang ist auch zu beachten, inwieweit Verfahren und Methoden neue Problemlagen und Herausforderungen für die Wissensproduktion wahrnehmen und adäquat repräsentieren können.

Dabei interessiert uns insbesondere der Zusammenhang zwischen externen Bewertungsprozessen und Mechanismen der internen Qualitätssicherung und -förderung vor allem im Hinblick darauf, wie interne Verfahren zur Vorbereitung externer

Evaluierungen dienen und damit besondere (bessere) Voraussetzungen für die Institutsakteure schaffen können und wie solche interne Verfahren wiederum in den Evaluationen wahrgenommen und bewertet werden. Hierzu erste Hinweise:

4.1 Sein oder Nicht-Sein: Die Evaluation durch den Wissenschaftsrat

(a) Die Vorbereitung des Instituts: Übung macht den Meister?

Deutlich wurde, daß in allen untersuchten Instituten die Wissenschaftsratsevaluation und ihre Vorbereitung sehr ernst genommen wurde; verbunden mit einem erheblichen Zeitaufwand und unter Einbeziehung verschiedenster Akteure - einschließlich der Personalvertretung und des wissenschaftlichen Beirats - sind intensive und strukturierte Institutsdebatten über das Forschungsprofil, die Forschungsfelder und ihre Weiterentwicklung sowie der institutionellen Rahmenbedingungen der Forschung initiiert worden; außerdem fanden Probebegehungen bzw. "Simulationen" der Begehung statt. In einigen Instituten wurden zusätzlich externe Gutachten erstellt oder externe Experten als Berater verpflichtet. In einem Fall wurden DFG-Antragsaktivitäten verstärkt - wohlwissend, worauf es ankommt.

Zwar nicht so intendiert oder als kontinuierlichen Prozeß der institutsinternen Selbstverständigung verstanden, ist die Vorbereitung auf die Evaluierung als produktiv empfunden worden, da sich mehrheitlich die wissenschaftlichen MitarbeiterInnen mit dem Institut und seinem Aufgaben- und Leistungsprofil auseinandersetzen und "über den Tellerrand" der eigenen Forschung blicken mußten. Nichtsdestotrotz wurde der damit verbundene extrem hohe Zeitaufwand problematisiert - vor allem im Hinblick auf die "normalen" Forschungsaufgaben im Institut, die in der Regel mit einem sehr knapp bemessenen Zeitbudget bewältigt werden müssen.

(b) Die Begehung: mehr Zeit, mehr Interaktion, weniger universitäre Sichtweisen

Im Sinne eines produktiven wissenschaftlichen Austausches mit den Evaluatoren empfanden die Institutsakteure die Möglichkeiten zur Darstellung der Forschungsarbeiten und zur Diskussion mit den Gutachtern im Rahmen der Begehung als nicht ausreichend, zudem der Grad der Vorbereitung und damit auch die Informiertheit der Gutachter als sehr unterschiedlich sowohl innerhalb der Kommissionen als auch

zwischen den Kommissionen für die jeweiligen Institute eingeschätzt wurde. Eine eindeutig positive Beurteilung erhielt die Diskussion der Wissenschaftsratsmitglieder mit den WissenschaftlerInnen unter Ausschluß des Leitungspersonals.

Neben der generellen Problematik, "unbefangene" und kompetente Gutachter für Evaluationen zu finden - und dies gerade angesichts der thematischen Breite und Vielfalt der Forschung der Blaue Liste-Institute - machte sich die universitär geprägte Sichtweise der Gutachter gerade bei einer stark anwendungsorientierten, politikbezogenen Forschung und den zum Teil sehr umfangreichen Service- und Beratungsleistungen der Institute als Problem bemerkbar. So wurde gegenüber einem Teil der abgegebenen Empfehlungen Unverständnis mit Blick auf die institutsspezifischen Aufgaben geäußert und eine Differenzierung als notwendig empfunden. Nutzer der Forschungsergebnisse sollten in den potentiellen Kreis von Evaluierern mit aufgenommen und zusätzlich eine Korrekturmöglichkeit von "Fehleinschätzungen" in den Evaluationsberichten eingeräumt werden.

(c) Die Auswirkungen: Anlaß zum Nachdenken

Im nachhinein der Evaluierung werden ihre positive Auswirkungen hervorgehoben: Konnten die Institute zwar nicht mit allen Empfehlungen des Wissenschaftsrats konform gehen und bestand aus Institutssicht auch nicht die Notwendigkeit, jede Empfehlung sofort umsetzen, so bewirkte der Prozeß der Evaluierung eine bessere Informiertheit, einen systematischen institutsinternen wissenschaftlichen Austausch über das Leistungsprofil und Entwicklungsoptionen, der für die Einrichtung nicht unbedingt eine Selbstverständlichkeit darstellt, eine bessere Transparenz, ein Problembewußtsein für abteilungsübergreifende Zusammenarbeit und eine bessere Kooperation zwischen den Forschungseinheiten. Defizite wurden institutsweit sichtbar und Diskussionsgegenstand: In dieser Form "nichtintendierte" Folgen hinsichtlich der Initiierung von institutionellen Lernprozessen in Forschungseinrichtungen, die kontinuierliche und systematische Diskussionen über den Status quo und Zukunftsperspektiven der Forschung nicht auf die Tagesordnung gesetzt haben.

Als ein erstes Fazit läßt sich daraus ziehen:

- Die Wissenschaftsratsevaluationen haben zum Teil erhebliche Defizite institutsinterner Kommunikations-, Selbstverständigungs- und Selbstkontrollprozesse im Sinne eines routinisierten Verfahrens zutage gebracht. Auch die Auswertung der Erfahrungen mit Evaluationsverfahren - sofern diese überhaupt stattgefunden hat - wurde nur punktuell auf Ziele, Methoden und Verfahren interner Selbstreflektionsprozesse bezogen, die durchaus auch Fragen der Selbststeuerung und Organisationsentwicklung von Forschungsinstituten tangieren.
- Trotz grundsätzlicher Akzeptanz von Evaluationen als integralem Bestandteil des Forschungs- und Wissenschaftssystems werden die Wissenschaftsratsevaluationen überwiegend als "Ausnahmestandard" hinsichtlich der Bewältigung von Institutsaufgaben begriffen; eine "Evaluationskultur" ist nicht identifizierbar. Institutspraktiken wie Simulationen von Begehungen des Wissenschaftsrats verweisen auf diesen Punkt - selbst wenn das Institut bereits mehrmals evaluiert wurde oder der wissenschaftliche Beirat explizite Evaluationsaufgaben übernimmt und so auf, wenn auch begrenzte, Routinisierungen rekuriert werden könnte.
- Da die Wissenschaftsratsevaluation nicht nur auf eine Bewertung der Forschungs- und Serviceleistungen der Blaue Liste-Institute abzielt, sondern auch Empfehlungen zur wissenschaftlichen Ausrichtung des Instituts, seiner institutionellen Voraussetzungen und der Rahmenbedingungen ausspricht, sollten Verfahrenselemente - wie die Institutsbegehung - hinsichtlich der Initiierung von Lernprozessen in den Instituten überdacht werden. Die beim Verfahren des Wissenschaftsrats einzukalkulierenden zeitlichen und personellen Restriktionen bei der Evaluierung von über 80 Blaue Liste-Institute spielen möglicherweise bei einem von den WGL-Gremien zu konzipierenden und durchzuführenden Verfahren eine nicht so bedeutende Rolle. Insbesondere sollte den Instituten nach Vorlage des Evaluationsberichts eine Feed-back-Runde mit den Evaluatoren ermöglicht werden, die im Sinne eines interaktiven Prozesses gestaltet werden könnte.

5. Evaluation als Qualitätssicherung und -förderung

Wie schon oben ausgeführt, findet zur Zeit vorrangig an den Hochschulen eine intensive Debatte über den Qualitätsbegriff vor allem bezüglich der Lehre, weniger der Forschung, sowie über entsprechende qualitätssichernde Maßnahmen und Verfahren statt. Für die außeruniversitären Forschungseinrichtungen und insbesondere die Blaue Liste-Institute ist der Diskurs hauptsächlich durch forschungspolitische Vorgaben von Akteuren des Wissenschaftssystems wie der Bund-Länder-Kommission, dem Wissenschaftsrat und dem Forschungsministerien bestimmt. Die Bund-Länder-Kommission definiert in ihrem Beschluß zur "Sicherheit der Qualität der Forschung" (siehe Abschnitt II.2.) zentrale Maßnahmen wie Peer Reviews, Programmbudgets und Kosten-Leistung-Rechnungen.

In den "Leitlinien" zur strategischen Orientierung der deutschen Forschungslandschaft, herausgegeben von dem ehemaligen Bundesminister für Bildung, Wissenschaft, Forschung und Technologie, Rüttgers, wird für Flexibilisierungsmöglichkeiten und verstärkten Wettbewerb plädiert und entsprechende Maßnahmen vorgeschlagen, die die Innovations- und damit die Leistungsfähigkeit der Forschung erhalten und fördern sollen. In dem Papier "Bildungs- und forschungspolitische Vorhaben und Schwerpunkte der Bundesregierung in der 14. Wahlperiode" der neuen Bundesministerin für Bildung und Forschung stehen die vorgestellten Schwerpunktsetzungen explizit unter dem Motto "Qualität hat Vorrang". Dafür soll ein wirksames Controlling-Instrument geschaffen werden als "Grundlage für mehr Transparenz, Effektivität, Effizienz und Qualität, zum Abbau von Verkrustung und Bürokratie". Fragen nach der Angemessenheit dieser Verfahren für die Forschungseinrichtungen hinsichtlich der damit verbundenen Ziele sowie die Frage, was ihre Qualität bzw. ihre Qualitäten eigentlich ausmacht, werden zur Zeit noch weitgehend ausgeblendet.

Wir wollten im Rahmen unserer Untersuchung erkunden, welche Vorstellungen seitens der Institute zu qualitätssichernden oder auch -fördernden Maßnahmen entwickelt wurden, ein Aspekt, dem nach wie vor zu wenig Aufmerksamkeit geschenkt wird. Diese Vorstellungen sind vor dem Hintergrund des spezifischen Aufgabenprofils und dem Verständnis der Institutsakteure von "guter" Forschung zu betrachten.

5.1 Qualitätssicherung (nicht nur) durch Evaluationen

Befragt nach den Zielen von Evaluationen außeruniversitärer Forschungseinrichtungen kommen neben allgemeinen Aspekten wie Legitimationsbedarf von öffentlich finanzierter Forschung, dem notwendigen Blick von außen als Korrektiv, Kontrolle und Förderung des wissenschaftlichen Niveaus, ja der Sichtweise von Evaluationen als unabdingbarem Bestandteil des Forschungssystems spezifische Anforderungen zum Ausdruck: Überprüft werden sollen die Besonderheiten der Institute, dies impliziert auch eine institutsadäquate Gestaltung des Verfahrens und beispielsweise die Einbeziehung von Nutzern der Forschung in den Evaluiererkreis. Erwartet wird mehrheitlich ein hoher Anteil an beratenden Funktionen und eine stärkere Rückkopplung zwischen Evaluatoren und dem Institut. Dementsprechend sollen Forschungsbewertungen eine kritische Reflexion der Standards bewirken und der Lokalisierung von Schwachstellen dienen. Hierfür ergibt sich die Notwendigkeit, Indikatoren für Fehlentwicklungen zu generieren. Eine entscheidende Voraussetzung für die produktive Gestaltung solcher Prozesse wird in dem Vertrauen der MitarbeiterInnen in die Institutsleitung gesehen. Hinsichtlich der Bedeutung des Forschungsmanagements und den entsprechenden Kompetenzen wird der Entwicklung von Elementen der Mitarbeiterführung in Evaluationsprozessen sowie der Evaluation von Personalpolitik Bedeutung geschenkt.

Die Einschätzung der Ziele und der Nutzen von institutionellen Selbstbeobachtungsverfahren verdeutlicht einerseits Skepsis der Institutsakteure gegenüber dem Nutzen von internen Evaluierungen, da diese Verfahren in der Regel hinsichtlich Anreize und Sanktionen folgenlos blieben, andererseits wird ihnen eine hohe Relevanz als Voraussetzung für Qualitätssicherung, Vorbereitung von externen Evaluierungen und als integraler Bestandteil des Forschungsprozesses beigemessen. Ein Institut führt ein auf quantitative Daten (insbesondere Drittmittelinwerbung, Publikationen in Refereed journals) basierendes Verfahren durch, das auf Wettbewerbsförderung zwischen den Abteilungen und einer - im begrenzten Umfang - leistungsabhängigen Allokation von Ressourcen ausgerichtet ist.

Eine gewisse Unsicherheit oder auch Unentschiedenheit der Akteure aus den Instituten gegenüber den Chancen und Problemen solcher Verfahren hängt möglicherweise auch mit dem Verweis auf andere Elemente der Praxis interner Qualitätssicherung

zusammen. Hingewiesen wird in diesem Zusammenhang auf zum Teil existierende interne Referee-Systeme, Publikationskommissionen, den wissenschaftlichen Austausch auf Fachkolloquien, Vorabpräsentation von Vorträgen vor einer institutsinternen Öffentlichkeit oder Beratungen bei der Drittmittelantragstellung; darüber hinaus zählen Verfahren der Forschungsplanung, der internen Steuerung und selbst Stellenbesetzungsverfahren zu qualitätssichernden Instrumenten.

Die Bedeutung der Zielsetzungen von internen Evaluierungen im Sinne einer kontinuierlichen Überprüfung der Leistungsfähigkeit, struktureller Rahmenbedingungen sowie der Identifizierung von Entwicklungsoptionen kommt u.a. in den Vorstellungen über die Aufgaben der wissenschaftlichen Institutsbeiräte zum Ausdruck: Sie sind primär als Beratungsinstanz in längerfristiger Sicht gefragt für strategische Orientierungen, bei der Entwicklung von Zukunftsperspektiven, bei der Identifizierung von Forschungstrends. Der Beirat soll durchaus "Finger in die offenen Wunden legen", und ihm wird als Instanz der intensiven Fachberatung ein hoher Stellenwert beigemessen. Mit Blick auf die vom Wissenschaftsrat zukünftig vorgesehenen Evaluierungsaufgaben für Beiräte problematisieren Institutsvertreter den zu erwartenden Spagat zwischen Beratung und Bewertung; zum Teil wird diese Rolle von Beiräten selber nicht akzeptiert. Ein Konzept für die Wahrnehmung beider Aufgaben scheint noch nicht in Sicht.

Die Frage nach der Zusammensetzung der Beiräte erfährt in allen Instituten eine hohe Relevanz. Neben einer notwendigen Aufgabenteilung zwischen Aufsichtsratsgremien, Mitgliederversammlungen (zum Teil sind die wissenschaftlichen Beiräte auch in diesen Gremien vertreten) und Beiräten soll die personelle Konstellation dem spezifischen Aufgabenprofil des Instituts entsprechen; dies impliziert beispielsweise die Repräsentanz der Industrie in diesen Gremien. Dort, wo bereits Vertreter von Wissenschaft und Forschung sowie der Industrie in den Beirat gewählt wurden, wird dieses Zusammenspiel als produktiv und nützlich für die Leistungsfähigkeit des Instituts empfunden.

Dieser Bedarf an Praxisvertretern in entsprechenden Institutsghremien muß vor dem Hintergrund der Forschungsausrichtung der Blaue Liste-Institute betrachtet werden. Wenn nicht der Anwendungsbezug der Forschung das Aufgabenprofil stark prägt, wird in anderen Fällen auf die notwendige Verbindung von Grundlagenforschung

und Problemorientierung, den gesellschaftlichen Problembezug hingewiesen. Befragt nach ihrem Verständnis von "guter Forschung", gehen die Institutsakteure mehrheitlich von der Forschung mit einem klaren Anwendungspotential, einer kurz- und langfristigen Problemlösungskapazität, einer Zweckgerichtetheit der Forschung und Aufgaben des Praxistransfers aus und betrachten die "saubere" Trennung von Grundlagen- und anwendungsorientierter Forschung als obsolet.

Drei Punkte möchten wir in diesem Zusammenhang hervorheben:

- Die von uns befragten Institute haben durchaus Verfahren der Qualitätssicherung und -förderung in den Forschungsprozeß integriert - sei es in Form von Präsentationen der Forschungsergebnisse, eines organisierten Prozesses der Forschungsplanung oder durch die Einrichtung von Publikationskommissionen zur Überprüfung der Qualität von Veröffentlichungen. Zum Teil werden diese strukturellen Elemente vom Institut nicht primär unter dem Aspekt der Qualitätssicherung wahrgenommen und nicht selber zum Gegenstand von institutsinternen Evaluierungen hinsichtlich ihrer Funktionalität, Reichweite (Kontrollen und Bewertungen) gemacht. Hier ist eine Chance zu sehen, bereits etablierte Strukturen kritisch zu reflektieren und weiter zu entwickeln und sie in interne Selbstkontroll- und Beobachtungsverfahren zu integrieren.
- Die mit Evaluationen verbundenen Erwartungen seitens der Institutsleitung und auch der wissenschaftlichen MitarbeiterInnen sind hoch: Eine systematische Differenzierung zwischen der Aufgabenstellung, den Akteuren und Adressaten interner Verfahren einerseits und externer Verfahren andererseits bzw. die gegenseitige Bezugnahme ist erst in Ansätzen entwickelt. Diese Frage ist auch auf die Entwicklung eines institutsangemessenen internen Indikatorensystems für Bewertungen zu beziehen.
- Die wissenschaftlichen Beiräte werden in Zukunft in ihrer Bedeutung wachsen, gerade wenn sie Evaluierungsaufgaben wahrnehmen. Ihre Beratungskompetenz wird auf Grundlage zum Teil dezidierter Institutskenntnisse hoch geschätzt und als ausbaufähig eingeschätzt. Um Beratungs- und Bewertungsaufgaben wahrnehmen zu können, bedarf es konzeptioneller Überlegungen zur (neuen) Rolle des Beirats und letztendlich auch explizierter Verfahrensregelungen mit Blick auf andere institutsinterne und auf externe wissenschaftspolitische Akteure.

V. Resümee

Im Blick auf die Zielbestimmung, Gegenstandsbereiche, Verfahrenselemente und Akteure von Forschungsevaluationen der Blaue Liste-Institute läßt sich ein erstes Resümee ziehen:

1. Die Evaluationen außeruniversitärer Forschungsinstitute (der Blauen Liste) richten sich auf die Bewertung der Forschung *und* ihre institutionellen Rahmenbedingungen. Hinsichtlich der Optimierung von Forschungsleistungen sind qualitätssichernde und -fördernde Maßnahmen als unabdingbare Bestandteile von Evaluationen zu betrachten. Angesichts der Vielfältigkeit und Unterschiedlichkeit der Aufgaben, Ziele, Organisationstypen, Wissensstrukturen und Bezugsgruppen der Institute müssen Evaluierungen ihre Spezifika erfassen. Im Zentrum steht die Frage, ob und wie die Institute ihre Ziele erreichen. Darüber hinaus sind die Ziele und Aufgabenbestimmungen selbst im Kontext innerwissenschaftlicher Entwicklungen und/oder (neuer) gesellschaftlicher Herausforderungen an die Forschung einer Überprüfung zu unterziehen.
2. Hinsichtlich des zu konzipierenden Evaluationsverfahrens für Blaue Liste-Institute sollte konzeptionell von einem integrierten und komplementären Modell ausgegangen werden, das interne und externe Selbstreflexions- und Begutachtungsprozesse beinhaltet. *Interne Evaluationen* im Sinne von regelmäßigen Selbstverständigungs- und Selbstreflexionsprozessen über den "Status quo" und das "Quo vadis" sind nicht nur unter dem vorbereitenden Aspekt externer Forschungsbewertungen zu betrachten, sondern haben ihre eigenständige Bedeutung als systematische Bestandsaufnahme, Leistungs- sowie Stärken- und Schwachstellenanalyse. Nicht auf "Einmaligkeit", sondern auf Regelmäßigkeit angelegt und auf die Erzeugung von Kommunikations- und Einigungsprozessen ausgerichtet, sind sie als Teil von Organisationsentwicklung zu begreifen. In diesem Sinne sollten die bereits eingeführten und etablierten Verfahren in Blaue Liste-Instituten zur Qualitätssicherung (beispielsweise Publication committees) sowie die internen Steuerungsgremien hinsichtlich ihrer Funktionalität von den Institutsakteuren überprüft werden.

Externe Evaluationen beziehen sich grundsätzlich auf dieselben Gegenstandsbe-
reiche und intendieren eine Überprüfung von Programmen, Zielen und Aufgaben-
stellungen sowie die Bereitstellung von Expertisen für qualitätsfördernde Maß-
nahmen. Sie unterscheiden sich jedoch insbesondere hinsichtlich des Grades der
Differenziertheit der zu erhebenden Daten von internen Evaluationen, da sie sich
an andere Adressaten richten. Die Existenz und Ausgestaltung von institutsinter-
nen Maßnahmen zur Qualitätskontrolle und -förderung stellen für externe Evaluie-
rungen ein entscheidendes Bewertungskriterium dar. Die Begutachtungen sind
komplementär zur Innensicht der jeweiligen Institute als externe "Kontrollinstanz"
zu verstehen.

3. Die Blaue Liste-Institute sind durch ein komplexes Aufgaben- und Forschungsfeld
- insbesondere die Verbindung von Grundlagen- und Anwendungsorientierung -
gekennzeichnet, das nicht nur durch die kognitive Struktur des Forschungsfeldes
bestimmt wird, sondern auch durch forschungsexterne Einflüsse. Neben den Lei-
stungserwartungen der scientific community, den Anforderungen der Zuwen-
dungsgeber und den Bedarfen von Anwendern und Praxisvertretern stehen die
Kriterien des Wissenschaftsrats zur Beurteilung der Forschung. Für die Komple-
xität der Forschung sowie für die Bearbeitung von interdisziplinären oder multi-
disziplinären Forschungsfeldern müssen institutionelle und forschungsstrukturelle
Lösungen gefunden werden. Als Herausforderung der betriebsförmig organisierten
Forschung stellen sich hier die Organisation von abteilungs- und insti-
tutsübergreifenden Kooperationsformen und von Dienstleistungen sowie die Ent-
wicklung von funktionalen internen Steuerungsmechanismen und -gremien dar.

Diese Komplexität von Aufgaben- und Leistungsprofilen sowie spezifischer Au-
ßenanforderungen und externer Kooperationsformen beeinflusst die Wahl der Eva-
luationsgegenstände: Neben den Forschungsergebnissen und der Forschungspro-
spektion stehen die Forschungsstruktur und -organisation hinsichtlich Funkiona-
lität und Adäquanz sowie die institutionellen Rahmenbedingungen (Ausstattung,
Finanzen etc.) für die Forschung zur Debatte.

4. Angesichts der Vielfältigkeit und Unterschiedlichkeit der Aufgaben der Blaue Liste-Institute gestaltet sich die "Meßbarkeit" der Qualität von Forschung besonders schwierig. Indikatoren müssen institutsadäquat das breite Spektrum unterschiedlicher Forschungstypen und der Dienst- und Serviceleistungen berücksichtigen. Die Voraussetzungen der Blaue Liste-Institute für die Bewertung durch Wissenschaftsindikatoren sind jedoch sehr unterschiedlich, da sie nur zum Teil grundlagenorientierte Forschungsfragen bearbeiten. Darüber hinaus werden Beratungsleistungen sowie Ergebnisse anwendungsorientierter Forschung bisher nicht adäquat erfaßt. Schließlich sind durch die eingesetzten Indikatoren erhebliche Rückwirkungen auf die Forschungspraxis in den Instituten identifizierbar. Indikatoren wirken als Steuerungsinstrumente, die auch unbeabsichtigte und negative Folgen haben können.

Um eine Bewertung der Blaue Liste-Institute nach vergleichbaren Maßstäben zu ermöglichen, ist bei der Entwicklung eines Indikatorensystems auf die breitestmögliche Erfassung des Aufgaben- und Leistungsspektrums der Blaue Liste-Institute besonderer Wert zu legen; darüber hinaus muß eine Flexibilität angestrebt werden, die zusätzlich eine Erfassung von Institutsspezifika erlaubt. Insbesondere bei der Entwicklung von Indikatorenprogrammen sollte letzter Punkt Beachtung finden.

5. Mit Blick auf die involvierten Akteure in den verschiedenen Evaluationskreisläufen wird den wissenschaftlichen Beiräten - als "Schnittstelle" zwischen internen Selbstverständigungsprozessen und externen Bewertungen - seitens der Institute eine besonders große Relevanz als kompetente und mit dem Institut vertraute Beratungsinstanz beigemessen: In mittel- und längerfristiger Sicht für strategische Orientierungen, bei der Entwicklung von Zukunftsperspektiven und der Identifizierung von Forschungstrends sowie bei Organisationsentscheidungen. Bei der vorgesehenen Übernahme von evaluierenden Aktivitäten durch Institutsbeiräte ist nicht von einer Unvereinbarkeit von Beratung und Bewertung auszugehen, wenn konzeptionell auf eine sorgsame Austarierung unterschiedlicher Aufgaben geachtet wird und Verfahrensregelungen entwickelt werden, die Beiräte nicht in unproduktive Rollenkonflikte mit dem Institut und externen wissenschaftspolitischen Akteuren bringen.

6. Ausgehend von der Orientierung der Evaluationen auf wissenschaftliche Bewertungen und Optimierungsstrategien, erfahren die Peers bei den externen Begutachtungen eine herausragende Bedeutung. Da der gesellschaftliche Problembezug oder direkte Anwendungsbezug der Forschung bei einem relevanten Teil der Institute das Aufgabenprofil entscheidend prägt, ist die Einbeziehung von kompetenten Praxisvertretern oder "Nutzern" der Forschung ein Desiderat.
7. Um die Akzeptanz der Bewertungen zu erhöhen und die Umsetzung der Empfehlungen zu befördern, sollte die *Vermittlung* der Ergebnisse externer Evaluationen - als Diskurs zwischen Evaluatoren und Akteuren der Institute - Bestandteil des Verfahrens sein.

Abkürzungen

AdW	Akademie der Wissenschaften der DDR
BMBF	Bundesministerium für Bildung und Forschung (Bis Oktober 1998 Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie)
BLK	Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung
DFG	Deutsche Forschungsgemeinschaft
FhG	Fraunhofer-Gesellschaft
HGF	Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren
HRK	Hochschulrektorenkonferenz
MPG	Max-Planck-Gesellschaft
WBL	Wissenschaftsgemeinschaft Blaue Liste
WGL	Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz
WZB	Wissenschaftszentrum Berlin für Sozialforschung

Literaturverzeichnis

Alewell, K., Autonomie mit Augenmaß, Göttingen 1993

Barz, A., Carstensen, D., Reissert R., Lehr- und Evaluationsberichte als Instrumente zur Qualitätsförderung. Bestandsaufnahme der aktuellen Praxis, Gütersloh 1997

Benz, W., Fruchtbare Kooperation, in: DUZ Extra, 17/1998

Blaschke, D., Zur Beurteilung interdisziplinärer sozialwissenschaftlicher Forschung, in: Fisch, R. / Daniel, H.-D. (Hg.), Messung und Förderung von Forschungsleistung. Person - Team - Institution, Konstanz 1986, S. 167 ff.

Block, H.-J. / Krull, W., What are the consequences? Reflections on the impact of evaluations conducted by a science policy advisory body, in: Scientometrics, 19/1990, S. 505 ff.

BLK (Hg.), Beschluß der Regierungschefs "Sicherung der Qualität der Forschung", Materialien zur Bildungsplanung und zur Forschungsförderung, Heft 61, Bonn 1998

BMBF (Hg.), Innovationen durch mehr Flexibilität und Wettbewerb. Leitlinien zur strategischen Orientierung der deutschen Forschungslandschaft, Bonn 1996

BMBF (Hg.), Faktenbericht 1998 zum Bundesbericht Forschung, Bonn 1998

Campbell, D. F. J. / Felderer, B., Evaluating Academic Research in Germany. Patterns and Policies, Vienna (Institute for Advanced Studies), Political Science Series No. 48, Vienna 1997

Daniel, Hans-Dieter und Rudolf Fisch (Hg.), Evaluation von Forschung. Methoden - Ergebnisse - Stellungnahmen, Konstanzer Beiträge zur sozialwissenschaftlichen Forschung Bd. 4, Konstanz 1988

- Daniel, H.- D., Guardians of Science. Fairness and Reliability of Peer Review, Weinheim u. a. 1993
- Evaluierungskommission der FhG, Systemevaluierung der Fraunhofer-Gesellschaft. Bericht der Evaluierungskommission, o. O., November 1998
- Etzioni, A., Soziologie der Organisation, München 1978
- Felt, U., Nowotny, H., Taschwer, K., Wissenschaftsforschung. Eine Einführung. Frankfurt/M. / New York 1995
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M., The new production of knowledge. The dynamics of science and research in contemporary societies, London and New Dehli 1994
- Goedegebuure, L. C. J., Maassen, P. A. M., Westerheijden, D. F., Peer Review and Performance Indicators: Quality Assessment in British and Dutch Higher Education, Utrecht 1990
- Gläser, J. / Meske, W., Anwendungsorientierung von Grundlagenforschung? Erfahrungen der Akademie der Wissenschaften der DDR, Frankfurt/M. / New York 1996
- Grundsatzausschuß der Arbeitsgemeinschaft Blaue Liste, Grundsätze und Arbeitsweisen der Wissenschaftsgemeinschaft Bund-Länder-Institute (WGBL), Entwurf vom 2.3.1995 (Manuskript)
- Grüner, H., Evaluation und Evaluationsforschung im Bildungswesen, in: Pädagogische Rundschau 47/1993, S. 29 ff.
- Hasse, R., Organisierte Forschung. Arbeitsteilung, Wettbewerb und Networking in Wissenschaft und Technik, Berlin 1996
- Hertel, I., Grußwort, in: Blümel, W., König, K. (Hg.), Verantwortung für die Forschung. Vorträge anlässlich des Festaktes zum 20jährigen Bestehen des Forschungsinstituts für öffentliche Verwaltung bei der Hochschule für Verwaltungswissenschaften Speyer am 23. September 1996, Berlin 1998, S. 41 ff.

- HGF (Hg.), Begutachtungen in der Helmholtz-Gemeinschaft. Grundsätze und Verfahren, Dokumentation 15, Bonn 1997
- Hochschulrektorenkonferenz (Hg.), Evaluation und Qualitätssicherung an den Hochschulen in Deutschland - Stand und Perspektiven, Beiträge zur Hochschulpolitik 6/1998, Bonn 1998
- Hochschulrektorenkonferenz (Hg.), Qualität an Hochschulen, Beiträge zur Hochschulpolitik 1/1999, Bonn 1999
- Hohn, H.-W., Kognitive Strukturen und Steuerungsprobleme der Forschung. Kernphysik und Informatik im Vergleich, Frankfurt/M. / New York 1998
- Hohn, H.-W., Schimank, U., Konflikte und Ungleichgewichte im Forschungssystem: Akteurkonstellationen und Entwicklungspfade in der außeruniversitären Forschung, Frankfurt/M. / New York 1990
- Hornbostel, S., Wissenschaftsindikatoren. Bewertungen in der Wissenschaft, Opladen 1997
- Krull, W., The Max Planck Experience of Evaluation, in: European Commission, First International Conference on the Evaluation of Research Technology and Development, 26, 27 and 28 April 1995, Thessaloniki, Greece, Scientometrics, Vol. 34 No. 3, S. 441 ff.
- Krull, W. / Winter, E. (Hg.), Reader zum Seminar "Von der Evaluation zur Prospektion? - Neue Herausforderungen für die deutsche Wissenschafts- und Forschungspolitik" (München, 10.-12. November 1995), Stifterverband für die Deutsche Wissenschaft, Essen 1995
- Lüthje, Jürgen, Evaluation im Verbund, in: Müller-Böling (Hg.), a. a. O., S. 148 ff.
- Maurer, M., Evaluation von Forschungsinstituten und -feldern durch den Wissenschaftsrat, in: Krull/Winter (Hg.), a.a.O. (ohne Seitenzahlen)
- MPG (Hg.), Die Max-Planck-Gesellschaft und ihre Institute, o. O., Dezember 1997

- Mayntz, R., Forschungsmanagement. Steuerungsversuche zwischen Scylla und Charybdis, Opladen 1985
- Mayntz, R. (Hg.), Deutsche Forschung im Einigungsprozeß. Die Transformation der Akademie der Wissenschaften der DDR 1989 bis 1992, Frankfurt/M. 1994
- Mittelstraß, J., Zukunft Forschung. Perspektiven der Hochschulforschung in einer Leonardo-Welt, in: Mittelstraß, J., Leonardo-Welt: über Wissenschaft, Forschung und Verantwortung, Frankfurt/M. 1992, S. 47 ff.
- Müller-Böling, D. (Hg.), Qualitätssicherung in Hochschulen. Forschung - Lehre - Management, Gütersloh 1995
- Neidhardt, F., Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG, Opladen 1988
- Neidhardt, F., Evaluationen im Hochschulbereich: Forschungsindikatoren, in: Gesellschaft von Freunden der Technischen Universität Berlin (Hg.), Humanismus und Technik, Jahrbuch 1995, Themenheft: Grundlagen und Probleme der Hochschulreform, 39. Band, Berlin 1996, S. 21 ff.
- Oppen, M., Qualitätsmanagement. Grundverständnisse, Umsetzungsstrategien und ein Erfolgsbericht: die Krankenkassen, Berlin 1995
- Röbbecke, M., Mitbestimmung und Forschungsorganisation, Baden-Baden 1997
- Seeling, S., Sproß, K., Wo bleiben die Hochschulen?, in: DUZ 4/1999, S. 8 ff.
- Simon, D., Im Block. Ein Vorsitzender blickt auf seine vielfältigen Erfahrungen zurück: Wie kluge Köpfe sich in raffinierten Kompromissen einzimmern, in: DIE ZEIT, Nr. 40 vom 26.09.97
- van Raan, A. F. J. (Hg.), Handbook of Quantitative Studies of Science and Technology, Amsterdam u.a. 1988

- Vroeijenstijn, A. I., Improvement and Accountability: Navigating between Scylla and Charybdis. Guide for External Quality Assessment in Higher Education, London/Bristol 1995
- Webler, W.-D., Qualitätssicherung in Lehre und Studium an deutschen Hochschulen, in: Zeitschrift für Sozialisationsforschung und Erziehungssoziologie, 1996/2, S. 119 ff.
- Webler, Wolff-Dietrich, Das Bielefelder Modell zur Evaluation der Lehre als Organisationsberatung durch Hochschulforscher, in: Evaluation und Qualitätssicherung an den Hochschulen in Deutschland - Stand und Perspektiven, Beiträge zur Hochschulpolitik 6/1998, Bonn 1998, S. 189 ff.
- Weingart, P., Prospektion und strategische Planung. Konzepte einer neuen gesellschaftsorientierten Wissenschaftspolitik, in: Wirtschaft und Wissenschaft, Jg. 3, 3/1995, S. 44 ff.
- Weingart, P., Sehringer, R., Winterhager, M. (Hg.), Indikatoren der Wissenschaft und Technik. Theorie, Methoden, Anwendungen, Frankfurt/M. / New York 1991
- Westerheijden, D. F., Brennan, J., Maassen, P. A. M., Changing Contexts of Quality Assessment. Recent Trends in West European Higher Education, Utrecht 1994
- WGBL (Hg.), Grundsätze und Arbeitsweisen der Wissenschaftsgemeinschaft Bund-Länder-Institute (WGBL), Vorschläge des Grundsatzausschusses der "Arbeitsgemeinschaft Blaue Liste" (AGBL) zur Realisierung der Neuordnung der Blauen Liste, Manuskript, o. O., 1995
- Wissenschaftsrat, Empfehlungen zur Neuordnung der Blauen Liste, Köln 1993
- Wissenschaftsrat, Stellungnahme zum Wissenschaftszentrum Berlin (WZB), Köln 1997
- Wolf, H.-G., Organisationsschicksale im deutschen Vereinigungsprozeß. Die Entwicklungswege der Institute der Akademie der Wissenschaften der DDR, Frankfurt/M./New York 1996